DOCUMENT RESUME

ED 217 541 EA 014 668

AUTHOR Cowden, Peter; Cohen, David K.
TITLE Divergent Worlds of Practice: The Federal Reform of Local Schools in the Experimental Schools Program. Final Report.
INSTITUTION Center for the Study of Public Policy, Cambridge, Mass.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE 81
CONTRACT 400-76-0128
NOTE 86p.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Educational Change; Educational Research; Elementary Secondary Education; Federal Government; *Federal Programs; *Government School Relationship; Program Evaluation; Research Utilization; *School District Autonomy
IDENTIFIERS *Experimental Schools Program

ABSTRACT

Focusing on the problems of the federal role in education, the educational "competency" of federal programs, and the usefulness of research-based knowledge for educational practice, this report evaluates the Experimental Schools Program (ESP), a small federal program in the early 1970s that aimed to help poor and minority students by encouraging "comprehensive change" in school districts' educational practices. Data for the evaluation came from ESP documents and interviews with ESP staffers, local school officials, and ESP researchers. After a general discussion of federal-local relations in education, the authors look at the differences between federal and local approaches to change, the variety of local responses to ESP, and the problem of defining comprehensive change. They next examine the conflicts arising from ESP's misunderstanding of school districts' decentralized structure and from its insistence on both local control and detailed federal monitoring. The following section describes ESP's holistic evaluations of its programs, the difficulties in applying holistic methods to the study of ill-defined changes, and the effects of evaluation on local administrators and teachers. The concluding section summarizes the authors' assessment of ESP. (RW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED217541

FINAL REPORT

DIVERGENT WORLDS OF PRACTICE

The Federal Reform of Local Schools in the Experimental Schools Program*

Peter Cowden and David K. Cohen

Cambridge, Massachusetts 02138

Summer 1981

* This study was supported by NIE through a grant to the Center for the Study of Public Policy (Contract No. 400-76-0128). Our thanks are due to Denis P. Doyle of NIE for his help and patience, and to all those in the Experimental Schools Program who gave generously of their time.

EA 014 668

I. INTRODUCTION

Arguments about the appropriate federal role in U.S. education are as old as public education itself, but the arguments became more intense as federal involvement in education grew. In the late 1950s and early 1960s these arguments centered on the political wisdom of federal intervention. In the late 1960s and 1970s, as evidence on the effects of the new Great Society education programs began to accumulate, the debate seemed to center more on whether federal intervention did any good, whatever one thought about its wisdom. But throughout these years, the federal role in education expanded, under both Democratic and Republican administrations. The 1980 presidential campaign once again placed the issue of political principle in the forefront, and since then the Reagan administration has begun trying to dismantle the results of two decades of growth, on the grounds that education should be chiefly a state and local responsibility.

Many issues intersect in these controversies, and disentangling them all would be no small task. One issue that has surfaced from time to time concerns the competence of the central government: what can it do well, or at least better than anyone else? Competence is of course not the only consideration in allocating political responsibilities, but it is not a trivial consideration. Competence can change with experience or resources, but there often are structural constraints on the development of competence, or great differences in the costs of developing it as one level of government rather than another. In this essay, we explore the competence of federal agencies in education, by exploring federal-local interaction in one education program.

One advantage of focusing on the competence of government agencies is that it directs attention to what the federal government has done well, and

poorly, in practice. Liberals have tended to assume that the government could do a good job of whatever it tried in education, if only the resources were available, while conservatives assume that a good job would be unlikely. Indeed, they believe that even if the central government could do a good job it would be a bad thing, because of the effects on the distribution of power within the federal system. Most of these arguments have occurred at the rarefied level of political principle. While we care about principle, the question of competence can best be illuminated by looking at practice. We consider the actual activities of federal and local educators, and the character of educational practice at the federal and local levels. Our story centers on the interactions among practitioners who work in these two very different parts of American education.

The case in point--the Experimental Schools Program--was modest in size, but broad in scope and ambition. ESP had a small total appropriation--of $55 million dollars--and supported projects in only 18 communities. But it was a highly touted effort by the U.S. Office of Education to bring "comprehensive" reform to local schools, and to create a new model of federal relations with local schools. ESP grew out of discontent with Great Society education policy and practice, and represented an effort to marry the sweeping visions of change associated with the Johnsonian programs to the concern for local control associated with more conservative views. The program began in 1971, but terminated in great disappointment and disarray a mere five years later.

Scrutiny of the Experimental Schools Program will

illuminate issues of federal competence, but it will not offer a conclusive answer to questions about the federal role. There is no such thing. Our account will throw only a bit of light on a complicated and continuing problem. We begin with a brief analysis of federal-local relations in education.

Federal-Local Relations

The federal role in education is ambiguous: since the late 1950s federal agencies have been charged increasingly with reforming education at the local level, but their resources for influencing local education are limited. The imbalance between federal mandates and federal resources has created many curious twists in federal-local relations. From a fiscal perspective, the relation seems plain enough: local schools receive a modest portion of their operating revenues from federal agencies, and federal education agencies impose various requirements for reform in return for the money. Typically, we think the local agencies are dependent upon their federal counterpart: the locals, after all, need the money. That is true enough, but there is a mutual obligation that implies a further dependence, one often unnoticed: for the federal agency's work to succeed, local schools must comply with federal requirements. If local school authorities depend on federal agencies for money, the federal agencies depend on local authorities to accomplish federal purposes.

The character of this dependence is complicated by several deep differences between federal and local educational practice. For one thing, the units of practice are different. Federal education managers have a national obligation, and their unit of practice is the federal program. They seek success for these programs, which means that to some extent they

must try to engage local districts in operations consistent with the program goals and values. They wish local districts to coordinate internal efforts toward those goals. The local units of practice, however, are schools and classrooms. There are tens of thousands of these in the United States--indeed, in any single large district there are thousands of classrooms. The numbers are so large, and the aims of education and the nature of teaching are so indeterminate, that even central administrators at the local level must leave teachers and principals a great deal of autonomy to shape their own practice. Education is coordinated at the local level, but in many crucial respects it is only weakly coordinated, at least by means of the formal local governmental and administrative apparatus. Coordination of local units of practice in federal programs is thus problematic from the outset.

In addition to differences in the units of practice, federal and local educators also differ concerning the aims of practice. Since the late 1950s, federal education managers have been preoccupied with the reform of local schools. Several factors contributed to the widespread assumption that local school reform is the central task of federal education agencies. One was external pressure on the federal government to promote change in local school policy and practice, because such change had been effectively blocked at the state or local level. The desegregation of southern schools is the best example of this phenomenon. A related factor is that Washington has become the nation's political center. The increasingly national character of politics and culture, and the widespread sense that the national government is the most powerful lever for social change, have encouraged those with an interest in reform of all sorts--even conservative reform--to gravitate to the national capital. And finally, there is the view, especially

common in education, that local government is inadequate, perhaps even bankrupt. These developments and others encouraged the belief in recent decades that the unique responsibility of federal policy in education is to correct the errors of local practice.

Reform does not figure so prominently in the work of local teachers and administrators. While many seek change in one way or another, their practice centers much more on providing complex services in an established social institution. Quite naturally, they are more preoccupied with what sociologists call maintenance than with change, for they must attend to all the continuing work of schools: they must select books, grade papers, teach subjects, keep schools heated, and hire and fire staff. Federal education officials, by contrast, have no schools to run, nor any of the associated tasks. They have federal programs to devise, administer, and change. Thus, not only do the units of practice differ greatly, but the nature of practitioners' work, and their responsibilities, differ as well. Powerful forces act upon local practice to press it away from an extensive focus on reform, while recent pressures on federal practice encourage precisely this focus.

Thus, the aims of federal practice have pressed federal agencies toward the reform of local practice, while differences in federal and local practice impede federal influence. These differences affect federal-local relations, but their effect varies according to the nature of the programs and policies over which relations occur. A federal program whose aim is simply to give financial support for whatever local education agencies wish to do--such as the Impact Aid Program--depends only modestly on local agencies

for success in its own performance; the Impact Aid Program is successful if LEAs simply accept the money and spend it on education. But a federal program that aims to reform local practice by spending money to improve students' school achievement--such as Title I of the 1965 ESEA--depends much more extensively upon local practice. In order for such a federal program to succeed, not only must money be spent by local educational authorities, but it must be spent on activities like reading, and spent in such a way that students' reading improves. That, in turn, means that there must also be changes in how teachers and other local practitioners do their work, and perhaps changes as well in the organization in which that work goes on.

In the case of the Title I program, then, the success of federal managers' practice depends on the success of many local practitioners' work. And the local work is complex, including not only teaching and learning reading, but managing those tasks, changing them, and quite likely also managing school organization. This is in marked contrast to unrestricted aid, where the success of federal practice would depend much less on the success of local practice. The chief requirement in the latter case would be local acceptance of federal monies and their successful expenditure on education: whether the activities on which the funds are expended were successful or not is probably irrelevant to whether the federal program has succeeded, or whether federal officials' work was competent. Since unrestricted aid would by definition not be targeted, and would be spent on many different local activities, federal responsibility for local practice would be greatly attenuated, and quite diffuse. There might be a sense in which the value of such aid would seem to be related to the quality of local education--i.e., if local education fell into wide disrepute, the aid might

be questioned. But in the case of programs like Title I, there would be a much greater and more specific federal dependence upon local practice: the program aims imply that the federal agency will improve the reading of local students, and the work of local practitioners in teaching and organizing reading. In some sense, federal officials become responsible for the work of local practitioners. Federal effectiveness must include the capacity to improve the effectiveness of their local partners. That is a tall order.

There is no reason in principle why such an order could not be filled, if federal agencies have the wherewithal--technical, political, or fiscal--to produce success in local schools. But if federal agencies lack the means to do so, then they will be seeking the reform of local practice without being able to muster the resources required to make the reforms a success. Federal agencies would then depend on local practitioners to make many complicated changes in their work, without being able to provide adequate means to induce or support those changes.

What resources do federal education agencies have at their disposal, to reform local practice? Money is one: if federal agencies paid much of the freight for local education--if they could increase or cut local budgets by large amounts--they might then have considerable influence on local practice. Political authority is another resource, for authority, however derived, would encourage local agencies to conform to federal mandates. A third resource is persuasion. Even if local schools have the dates. A third resource is persuasion. For even if local schools have the other resources they need to carry out federal mandates, they may lack the desire to do so. Or, having the desire, they may lack the know-how. Federal agencies might remedy these problems by persuasion. They could reason with

local educators, convincing them by rational argument or appeals to morality that federal mandates should be carried out. Or federal officials could remedy the lack of local know-how by supplying it to the deficient localities, or by providing inducements for the localities to acquire it themselves, or both.[1]

Federal education agencies have used all of these resources in their efforts to secure local compliance with federal policy in education. One difficulty in their efforts, though, has been the considerable limits on federal authority in education: by law and tradition, most authority in education lies at the state and local level. Authority is an extremely limited federal resource. Only in a few extraordinary situations has the federal government commanded the moral or political resources to exercise authority effectively. Another difficulty has been that federal monies, while much increased over earlier years, are still a modest fraction of local revenues. Local districts value federal funds because they offer room for maneuver at the margin, and this provides federal education agencies with some leverage on localities. But the leverage has been limited, in part by the size of the federal contribution. In these two respects, then, federal resources for influencing local action have been only modest.

As a result, federal education agencies have been chronically overextended: the scope of their mandates for local change typically far exceeded the money and authority the central government could bring to bear on state and local school agencies. The only exception to this pattern, briefly, was Title VI of the 1964 Civil Rights Act. For a few years during the middle and late 1960s, federal officials had something so powerful to offer local practitioners--namely, the absence of fiscal, political, and legal trouble,

perhaps even the absence of civil strife--that local officials were encouraged, often against their will, to comply with federal directives. As a result, many southern schools were desegregated.

But in most cases, federal officials have had less to offer: statutory requirements for participation; some ideas about how to do it; regulations; and a modest amount of money. These instruments are sufficient to produce the local changes federal officials desired when local practitioners both wished a particular change and had the capacity to produce it. This was the case in many instances, and local practice changed. In many other instances, however, the locals did not wish the changes federal agencies desired, or did not wish all the federally mandated changes, or did not wish to make the changes in the ways required, or could not make the changes despite their wish to do so. In such cases, federal agencies depended on local schools to reform local practice, when either the local will to make the changes, or the local ability to change, or both were in question. Federal agencies were asserting national priorities, but because they lacked the resources to insure local compliance with these priorities, the success of federal programs came to depend upon local agencies whose capacity for change was doubtful. That is a curious sort of dependence.

One result of this situation was a peculiar pattern in federal-local political relations. Because local educators could see the imbalance between federal mandates and federal resources, they knew that in fact they had considerable discretion in carrying out those mandates, even when the regulations and guidelines might suggest otherwise. From this angle, Title I of the 1965 ESEA, and PL 94-142, which were treated in regulations and guidelines as though they were highly restricted categorical

programs, might be thought of as being moderately restricted block grants in practice.

Such incongruence is not necessarily a problem, if everyone understands the limits it implies, and observes a decent respect for the proprieties. But the federal preoccupation with reform often led to a curious interpretation of events: because many federal officials saw their task as the improvement of local practice that was sadly deficient, they were inclined to interpret local behavior in terms of this presumed deficiency. Federal officials tended to explain poor local performance in federal programs as another manifestation of local deficiency. Other explanations, such as overambitious program aims, or insufficient program resources, were frequently ignored. As a result, in many cases when federal programs appeared to produce poor results, federal executives and many other observers concluded that the problem lay at the local level, and that the solution was a stronger dose of the previous medicine. This approach led to more ideas for reform of local practice, and to more statutes, regulations, and guidelines. The imbalance between federal requirements and federal resources was not corrected, though, so that as federal requirements for local reform mounted, federal resources remained only a modest contribution to the local costs of reform. The ensuing imbalance insured that many federal ambitions would fail of achievement, and would themselves be seen as evidence of the need for even more ambitious federal action. The Experimental Schools Program is a particularly nice case of this crazy logic, in which federal reformers seemed incapable of imagining any other reason for their program's problems than local incapacity.

A second result of the dependence of federal reformers on local agencies is

a need to rely on other than fiscal and political resources as a means of accomplishing federal objectives, namely on persuasion. Persuasion has many forms--including everything from the use of rational argument to decide about a program of research, to brainwashing an entire people. In the case of federal education programs, the particular form of persuasion that federal authorities relied on was social science. One hope behind the federal investment in social science was that research could identify the solutions to educational problems. Another hope was that if researchers could identify better teaching methods, or produce better curricula, the knowledge itself would stimulate local action. Consequently, as federal officials pressed for local school reform they made extraordinary efforts to develop social science knowledge about what good educational practice is, and how it can be promoted. Since the early 1960s, when federal education officials have assessed the schools' shortcomings they have identified the absence of knowledge about good practice and how to produce it as a key problem. And repeated efforts have been made, in one program after another, to remedy this problem by more and better research and evaluation. This pattern has been evident in nearly everything from early efforts to establish R&D Centers and Regional Educational Laboratories designed to develop new knowledge that would improve local practice, to recent attempts to improve evaluation in such programs as Title I ESEA, by developing canned "models" for evaluation and supporting large training and technical assistance agencies to help local schools to use these models to improve Title I evaluation. Knowledge, after all, is a resource that federal practitioners might well use in large amounts, even if money and political power were in short supply. What federal education agencies lacked in fiscal and political clout, they tried to make up with scientific knowledge.

Federal officials turned to applied social science on the assumption that it could be authoritative in the quest for improved local practice. In contrast to the patchy, anecdotal, often inarticulate and hand-me-down knowledge of local practitioners, social science was believed to be capable of demonstrating, in a scientifically conclusive way, which local practices worked, and why. Because it would be conclusive, social science was expected to provide compelling inducements for action to improve schools.[2] But as it turned out, social science was not conclusive about what programs worked or didn't: doing more evaluation did not decrease scientific disagreement about what works, or what we mean by "works," or why some things work and others don't. If anything, evaluations and other studies increased controversy on these points. Social science did not establish an authoritative basis for improving practice, and thus did not provide federal education agencies with a reliable resource for reforming local practice.

Curiously enough, though, regularly poor results with evaluation did not discourage federal reliance on social science as an instrument of policy. Instead, the failures appear to have had the opposite effect: they were taken as a sign that more applied research was needed, in order to produce the desired knowledge, and thereby the reforms of local practice.

The history of federal efforts to reform local practice was thus marked by more and more disappointing results from social science, yet at the same time, by ever more projects for local reform through social science, and more money with which to do the job. The response to the failure of social science was demands for more and better social science. One important consequence was a rapid growth of federal research and development efforts into many corners of education. Another was an extraordinarily broad definition of the tasks of research and development. As the poor results of some applied

research become apparent, federal administrators expanded the tasks of research even further, in a search for ways to compensate for the unexpected failures. The response, for example, to failures of early program evaluation was a flurry of attempts to scientifically distinguish successful from unsuccessful local projects, to determine what produced success in the happy cases and how it could be replicated. The original scientific agenda--determining whether a particular program succeeded--gave way, upon its failure, to an expanded agenda: finding projects that work, figuring out why some projects succeed and other fail, and figuring out how to get complex social institutions to replicate success. There were many instances of this phenomenon. The wish to understand social problems is reasonable, but it was perverted by a crazy logic that inferred, from the failure of already ambitious research, a need for even more ambitious research.

One result of that logic was to debase the modest jobs that social research can do well, by setting unattainable tasks that social research cannot perform. Another result was to needlessly magnify the failures of applied research; by the end of the 1970s, educational R&D had produced unimpressive results all across a large menu of impossible tasks. This did nothing to improve the reputation, either of social science or of its applications.

Still another result of the growth of applied research in education--or perhaps a cause of it--was a tendency to confuse scientific analyses of practice with what teachers and administrators actually do. The odd idea grew up that through social science, federal administrators can know enough about educational practice to make great changes in the work of local teachers and administrators. This notion lay at the heart of the Experimental Schools Program, among many other activities in the last two decades. Federal managers believed that applied research could actually reform the knowledge that local teachers and administrators used in their work. ESP

managers, for example, believed that if they devised a new and more comprehensive scheme for social science evaluation, reliable scientific information could replace the knowledge commonly used in managing school districts, local schools, and classrooms. They ignored the frailties of social science, and ignored as well the strengths of both ordinary knowledge and the accumulated professional knowledge of teachers and administrators. Like many other federal reformers of the last two decades, ESP managers assumed that social science information can provide authoritative guides to action, and that these can strongly influence what teachers and administrators do.

But social science, while useful for many things, could not bear the weight that federal administrators tried to place upon it. The ESP evaluation was a resounding and costly failure. It was instructive, though, for it encourages us to to ask whether, among practitioners' many resources in education, scientific knowledge can play more than a very modest role. However excellent the studies, can educational practice depend heavily on social science? Can government use social science as an effective instrument of persuasion? Scrutiny of the Experimental Schools Program can illuminate these questions, among others.

II. FEDERAL VS. LOCAL APPROACHES TO CHANGE

The aim of ESP was comprehensive change. This was seen as a new way of thinking about how to change schools, in contrast to other federal programs that made what were viewed as "piecemeal" or "categorical" change (e.g., Title I).[1] The notion of comprehensive change is familiar in federal social policy,[2] but the federal ESP staff saw it as a novel approach to reform in education. An internal ESP memo explained that piecemeal efforts, such as "curriculum reform projects, failed because the

[new] curriculum presupposed knowledge or skills not present in teachers who were not given adequate training and the curriculum as intended never was implemented."[3] The thrust of the criticism was that there was little coordination between two elements of a school system, teachers and the curriculum, and so reforms failed. Another memo explained that new programs "were introduced without adequate involvement and/or preparation of a community which raised the normal level of anxiety associated with any change to the point where any problem encountered in the implementation of the program could be sufficient to kill it."[4] The implication was that such programs had failed because they did not include one of the components considered essential to coherent reform: community participation in the local education agency (LEA).

The ESP staff thus believed that significant improvements in education were unlikely without "holistic" or "comprehensive" change, i.e., unless the different "practices" or separate "pieces" were compatible and mutually reinforcing in a "synergistic" way.[5] They further assumed that a school system can be neatly described, and made to work coherently. The ESP description explained that change must "include but not be limited to" the "five major" parts of a school system: "curriculum development, community participation, staff development, administration, and organization";[6] if there was coordination across these five parts of the system, coherent change would result. Based on this assumption, the federal ESP staff required school districts to define their plan for change via an "educational concept" or "central theme" representing the district's plans to get from "what exists at present to what education ought to be in terms of its needs and aspirations."[7]

The Local Response

Local approaches to the program differed substantially from federal reformers' hopes. As one federal monitor reported, local sites, especially those in rural areas, could not grasp the meaning of comprehensive change. In response, federal monitors often treated local personnel "like graduate students who had a dissertation to defend."[8] Another federal monitor touched on the heart of the difference in federal and local conceptions of change when she said that local projects "refused to think of pieces as interlocking parts."[9] Local practice simply did not reflect the coherence assumed in the federal notion of comprehensive change.

Problems at the Administrative Level

Local administrators typically interpreted comprehensive change in terms of local needs and operations. One superintendent involved in curriculum reform explained that he was attracted to the program because it "looked broad enough so that we could attack any curriculum problem we could find." He related comprehensiveness only to curriculum; and he chose curriculum from some sense of its compatibility with current work in the school district. He went on to say that "what is comprehensive in math is that it is K-12, what is comprehensive in history is that it is community-wide in developing (sic)." He was perplexed and perturbed, like others in his position, when ESP Washington rejected his proposed program. He summed up the problem by saying that the Washington program director was "hard on comprehensive change . . . and no matter what we told him it was never comprehensive . . . . He did not want us to just buy text-

books, he wanted us to think bigger." The superintendent was never sure what "bigger" was.[10] At the local level, changing the curriculum seemed a sweeping and "comprehensive" reform program. From the federal perspective, however, change in all five areas, not just curriculum change, seemed essential.

Local administrators in another site ran into similar problems. Asked what he understood by comprehensive change, the superintendent replied that he "blocked on the word" because he had "spent so many months at it." He added that the "feds" kept demanding that the district define comprehensiveness, but that "the game [the district] got into was a series of Title III projects."[11] These were piecemeal educational innovations such as individualized curricula or supplementary service centers. They reflected little joint decision making, for the local project directors confined themselves to their own projects. This was precisely what federal reformers did not intend; instead of a single coherent program of district-wide reform, they found ESP supporting piecemeal local projects, running independently of one another without common focus.

In another site, the superintendent was preoccupied with expanding existing federally funded local projects. In proposals to ESP, he sought to connect a number of these by explaining that "comprehensiveness" lay in the "optimum learning environment" of the school.[12] But the connection between projects was defined only vaguely and retrospectively, and involved little decision making across projects. His newsletter elaborated on his interpretation of comprehensiveness: the "proposed [ESP] project could mean that many federal projects now operating . . . would be expanded into an

operational district-wide setup. The stress is not on developing a new program but upon operating the many programs now established throughout the district."[13]

Once again, then, the local definition of comprehensive change rested on what was already occurring in the school district. Like most definitions offered, it was confusing. One reason for this is that the concept itself is indistinct. Another is that local administrators attempted to unify unrelated activities by means of words designed to satsify federal monitors. But if the words portrayed a broad program of integrated change, the local projects actually consisted of incremental modifications of existing projects little related to one another. One superintendent discussed his view of ESP Washington's negative reaction: he felt that the project director of the National Institute of Education (NIE) "harped" on the fact that his proposed program was not "comprehensive."[14] He saw the expansion of activities across the district as promising comprehensive change because it touched on so many areas of practice, disjointed though they were. The federal staff, by contrast, saw his program as an attempt to make disconnected projects appear to be a coherent whole.

Conflict over the meaning of comprehensive change was thus a persistent feature of the program. Local administrators continually proposed plans to incrementally modify varied and often unrelated existing endeavors, while the federal staff wanted all endeavors to reflect an overall, consistent, district-wide plan, and a big change in the way of doing business. Having won ESP grants primarily because they had shown a willingness to make comprehensive change, local sites could produce no specific plans to that end. For one thing, local operations themselves had not the coherence that would support such a plan. Moreover, local managers

were unsure what comprehensive change was. They proposed and reproposed programs of change; the federal ESP rejected them as lacking the coherence required for comprehensive change. As one superintendent reported, after rewriting his district plan several times he "did not know whether [it] was finally right or not, but I guess it was because they approved it."[15] Another project director, asked about comprehensiveness, explained, "We never did resolve that damned thing, we never had any agreement on what it meant."[16] In at least one case, a local site refused to rewrite its plans after several attempts.[17] As a last resort, the federal ESP staff often just approved the most recent plan, hoping that local understanding of a comprehensive approach might grow as the program proceeded. The program had to be maintained, even if local projects were wanting.

The difference in approaches to change was also manifest in local sites' inability to offer acceptable central themes for their projects. The federal staff intended that theme to be a common basis for changes in different parts of the school system, in conformity with their assumption that comprehensiveness would assure a new coherence in local practice. Yet a central theme consistently eluded local practitioners. Administrators saw potential change occurring not in accordance with any single theme, but rather in various ways in different parts of the school system. One local project director, in explaining federal rejection of his plan, confessed that he "did not know what [the program] was all about, except there was supposed to be a theme for the project." He went on, "Every time we came up with something, [the federal monitor] did not think it was what it was supposed to be. And I just did not know what she wanted. Every time I

submitted a theme, one sentence mind you, she rejected it."[18]

The federal ESP staff saw proposed themes as either too vague (e.g., "individualization") or so specific (e.g., curriculum reform) that they ran counter to the notion of coherent change involving the five elements of the school system. To local administrators, the requirement for a central theme was a contrived rationale for different programs and activities only tenuously related to one another--an academic exercise. To the ESP Washington staff, the theme was to indicate the main direction of a district-wide program of change.

Ironically, local incoherence was to some extent fostered by federal managers. Local administrators used ESP to support various specific efforts at change, because they were pressed to support many different activities and projects in their districts: typically there were more local requests for project support than could be supported. Yet federal ESP staff kept exhorting local administrators toward broad community involvement in the change process, because they believed that this would promote coherence. Federal officials did not seem to realize that, given the considerable autonomy of sub-units in local school systems, the greater the number of local participants involved the more disparate and incoherent the program would become. One superintendent commented: "The [ESP staff in Washington] kept saying that we were not opening it up enough, but we thought we had opened it up too much."[19] For the local administration it was a Catch-22 predicament: the more accessible ESP was made to the district, the more diverse and numerous were the requests by those seeking support for their particular pro-

grams. Thus as local administrators complied with federal demands for greater local participation, the program was increasingly shaped by local groups' diverse interests.

Problems at the School Level

Some of the reasons why local sites did not perform up to federal expectations can be seen in principals' and and teachers' responses to ESP. Principals within school districts had diverse conceptions of ESP, only loosely connected to those of central administrators. In one site, where the central administration's "theme" was humanism and individualization, a principal explained: "We kept getting the words of humanism and individualization, but I think nobody had the same meaning about what we meant."[20] A principal in another district explained that for him or anyone else, the central administration's notion of the program was "hard to say in a detailed sense without lots of false statements and misunderstandings."[21] Another principal recalled, "It soon became obvious to me that [the local project director] did not know what was going on and the other principals did not know either. So he would keep asking us [what our ESP project was about] because he did not know."[22]

The diverse ways in which principals within a school district thought about change were due partly to the fact that different changes were desired in different schools. Operating largely independently of one another, principals had little reason to want to understand, let alone to rely upon other principals' thoughts about ESP. Their interest was limited to what was feasible in their school. As one principal remarked, "At first I thought the project was

offering a lot of work I did not need. But then I realized that I could do lots of things I did not see before."[23] Another principal said that he saw the program "as an opportunity to develop different programs we could not get from other federal agencies."[24] Typically, local change was seen only in terms of a principal's own school and was thus inconsistent with federal notions of district-wide "synergistic" efforts at "comprehensive change."

Teachers' reactions were similar. Often they didn't even know about comprehensive change, despite the efforts of some principals and central administrators to explain the program. One senior federal monitor, after visiting a highly regarded project after a year of operation, reported disdainfully, "My speech to the teachers regarding the central theme of experimental schools and of the local concerns clearly was the first time they had heard of comprehensiveness."[25] In addition, teachers' views of ESP generally had little to do with what central administrators or principals had in mind. The program often was of no interest at all. One teacher commented, "I'm not into this alternative thing [the 'theme' of her site]; that is their bag. I am into the classroom and do my own thing there."[26] Another was "sure good things came out of it," but could not identify what they might be.[27] And among those teachers who thought they were working in ESP--many didn't even know it-- each one's activities seemed little related to what other teachers in the school were doing. As one teacher explained, "From what observation I have made, I have seen as much variety going on in one school as there is in any single program [of school-

wide change]."[28] But since ultimately the change took place at the classroom level, in accordance with each teacher's interpretation, even school-wide change often proved illusory. According to one ESP project staff member, "Everybody worked like hell--like horses pulling in different directions."[29]

Some inertia was certainly involved. Teachers wanted to continue what they were doing before ESP, and they did. As one teacher put it, "For the most part, everybody protected the status quo . . . after the initial thrust everybody protected what they had."[30] Others were hostile or simply apathetic toward change efforts initiated by central administrators or principals. One teacher described the feeling in her school as wanting to be left alone, and was herself "downright negative" to ESP.[31] A curriculum coordinator explained, in rather typical fashion, what happened when the new curriculum was brought into her school: "Some of the kids were going to get a self-concept and self-awareness (sic) so they brought all sorts of material but little came of it."[32] A classroom teacher said sarcastically that of several new curriculum reforms tried in her school, "The one that had the most success and made the biggest single contribution was the Pullman Manual for Physical Education."[33] She meant that while only a few teachers used the new curriculum, the Pullman Manual succeeded because there was only one physical education teacher, who used it.

Problems at the Federal Level

The unexpected variety of images and interpretations evoked by ESP at the local level was troublesome for federal managers. Some

ESP staffers couldn't figure out whether various activities in a local site were related, or "comprehensive." Nor did federal reformers agree among themselves whether what was occurring in a site was significant or not--or in some cases whether anything was happening at all. After a site visit, one federal monitor commented upon another monitor's perceptions of a prized site, believed to be successfully making significant comprehensive change. Her memo explained that, contrary to her fellow monitor's perceptions, "the local project certainly was not comprehensive and, further, no one involved in writing the plan, executing it, administering it, watching it, was aware of what comprehensive change meant."[34] She concluded that in this site ESP had produced little change that was not occurring before the program.

Conflict and confusion were not limited to the staff. Even the first director of NIE, after assuming responsibility for ESP from the U.S. Office of Education,[35] was unclear about the program and the meaning of comprehensive change. In an interview he stated that he had never understood "what Binswanger [the ESP director] was getting at with his notion of comprehensive change." It struck him as "more an experiment in government funding and intervention techniques rather than an experiment in educational concepts and techniques." He further explained that the program had already failed when he took it over; and that it continued because, as with other programs of reform, terminating it was a "political impossibility."[36]

Nonetheless, ESP program staffers held to the view that

ESP failed for the same reason earlier reform efforts had failed: local resistance to federal plans and intentions. Local sites, they held, were unwilling to change as their central administrations had at first promised. Instead, according to one federal monitor, local sites had "done what [they] needed"--that is, used the program to meet local needs.[37] Others saw local sites simply as "ripping off" the government, or engaging in activities that were "fraudulent."[38] As one senior federal monitor said, reflecting on the program, "I would do it again but with [local] people who believed in it."[39]

Footnotes to Chapter II*

[1]The interest in comprehensiveness was echoed in a number of reports. See, for example, The National Advisory Council on the Education of Disadvantaged Children. Annual Report, December 5, 1966. The report talks about how "Title I projects are piecemeal fragmented efforts at remediation or vaguely directed 'enrichment.'" (Quoted in Berke, J. S. & Kirst, M. W. Federal Aid to Education: Who Benefits? Lexington, MA.: D.C. Heath and Company, 1972, p. 396.)

[2]See, for example, Rein, M. The coordination of social sciences. In M. Rein, Social Policy. New York: Random House, 1970; Aiken, M., et al. Coordinating Human Services. San Francisco: Jossey Bass, 1975; Cohen, D. Coordination in Social Services (Monograph). Cambridge, MA.: Harvard University Graduate School of Education, Center for Educational Policy Research, June 1974.

[3]Internal memo from the files of Experimental Schools Program [hereafter, ESP] office in Washington, undated and anonymous.

[4]Ibid.

[5]Binswanger, R. Testimony before the House Appropriations Subcommittee on Education, 92nd Congress, 1st Session, March 3, 1971. Congressional Hearings (Vol. 63), p. 938.

[6]Department of Health, Education, and Welfare document, Basic Program Information, nationally distributed to local school districts, dated December 28, 1970.

[7]Ibid.

[8]Interview with former local ESP project staff member and former federal ESP staff member, March 31, 1977.

[9]Interview with former senior federal ESP staff member, November 17, 1976.

[10]Interview with district superintendent of a local rural school site (anonymity requested), May 2, 1977.

[11]Interview with local ESP project director in the South Umqua School District, South Umqua, Oregon, July 8, 1977.

---

*An effort has been made to provide for the anonymity of the respondents interviewed in this study by omitting their names. Interviewees are identified only by role and location, unless they requested otherwise.

[12]Interview with district superintendent of the Franklin Pierce School District, Tacoma, Washington, March 30, 1977.

[13]Internal staff newsletter of the Franklin Pierce School District, Tacoma, Washington, dated February 1971.

[14]Interview with district superintendent of the Franklin Pierce School District, Tacoma, Washington, March 30, 1977.

[15]Interview with district superintendent of a rural school site (site anonymity requested), May 2, 1977.

[16]Interview with local ESP project director in the Berkeley Unified School District, Berkeley, California, May 10, 1977.

[17]Interview with district superintendent of the Carbon County School District, Saratoga, Wyoming, April 21, 1977.

[18]Interview with local ESP project director in the Carbon County School District, Saratoga, Wyoming, April 21, 1977.

[19]Interview with district superintendent of the South Umqua School District, South Umqua, Oregon, May 23, 1977.

[20]Interview with school principal in the Constantine School District, Constantine, Michigan, May 3, 1977.

[21]Interview with school principal in Supervisory Union #58, Groveton, New Hampshire, June 10, 1977.

[22]Interview with school principal at the Encampment School, Saratoga, Wyoming, April 22, 1977.

[23]Interview with school principal in the Franklin Pierce School District, Tacoma, Washington, March 30, 1977.

[24]Interview with school principal in Supervisory Union #58, Groveton, New Hampshire, June 10, 1977.

[25]Personal memo to the files on the history of ESP by a senior federal ESP staff member, dated December 30, 1973.

[26]Interview with teacher in the Southeast Alternative School District, Minneapolis, Minnesota, June 2, 1977.

[27]Interview with teacher in the South Umqua School District, South Umqua, Oregon, May 24, 1977.

[28]Interview with English department chairman in the Southeast Alternative School District, Minneapolis, Minnesota, June 1, 1977.

[29]Interview with local ESP project staff member in the Edgewood Independent School District, San Antonio, Texas, April 12, 1977.

[30]Interview with teacher in the Southeast Alternative School District, Minneapolis, Minnesota, May 31, 1977.

[31]Interview with teacher in Supervisory Union #58, Groveton, New Hampshire, June 9, 1977.

[32]Interview with school curriculum arts coordinator in Supervisory Union #58, Groveton, New Hampshire, June 9, 1977.

[33]Interview with teacher in the Franklin Pierce School District, Tacoma, Washington, March 30, 1977.

[34]Personal memo to the files on the history of ESP by a senior federal ESP staff member, dated December 30, 1973.

[35]ESP was transferred from the Office of Education to the newly created National Institute of Education (NIE) a year and a half after its inception. For a more complete discussion of the effects of the creation of NIE on programs, see Sproull, L., et al. Organizing Anarchy: Belief, Bureaucracy and Politics in the National Institute of Education. Chicago: University of Chicago Press, 1978.

[36]Interview with former director of the National Institute of Education, October 24, 1976.

[37]Interview with former senior federal ESP staff member, November 17, 1976.

[38]Interview with former federal ESP staff member, December 3, 1976.

[39]Interview with former senior federal ESP staff member, November 17, 1976.

III. FEDERAL VS. LOCAL UNITS OF PRACTICE

One essential part of the ESP idea was to insure local commitment by insuring local control. The federal managers assumed that if local administrators felt that the program was their own, they would do a better job. The federal ESP staff saw their primary task as giving control to local administrators, and holding them accountable for implementing the programs. This, of course, assumed that comprehensive change in a district could be managed in a centralized way--that control lay in the central administration.[1] This idea made a certain kind of sense for federal practitioners, for the unit of their practice was the program: how could they deal with each local school individually? But the control of local central administrators is limited: the units of local practice are schools and classrooms, which are at least semi-autonomous. Local administrators cannot manage school districts without permitting a great deal of autonomy for teachers and principals. In managing the ESP program, they followed this pattern, delegating control to the school level, and acquiescing in principals' and teachers' plans for change. In the decentralized organization of local practice, the operations needed to insure the interest and participation of the local units conflicted with the federal assumption that district-wide change could be centrally managed.

Local Control as Federal Policy

At the time of ESP's inception, fostering local autonomy was encouraged in federal policy circles. The previous decade's experience had convinced many federal reformers that simply "laying on money in the name of doing good" did not solve social problems.[2] Moreover, local educators had been complaining about the insensitivity and domination

of the federal bureaucracy. Federal reformers were accused of alienating local school practitioners and suffocating local commitment in their programs for educational change. The federal director of ESP, Robert Binswanger, was particularly aware of these concerns, and committed to local control. One of his colleagues explained that it was this aspect of ESP which "turned [Binswanger] on." Binswanger wanted "to have a program which was really responsive to local schools--to give a good idea a try--not that it would necessarily work, but to try, and he would protect it from the feds."[3] Tom Glennan, the first director of NIE, explained that ESP had emerged from a Congressional climate in which federal control of local programs was being severely criticized: "Schools should be able to do what they want."[4] Another member of NIE, formerly on the ESP staff in Washington, said that ESP was conceived at a time when any concept that included local control was popular.[5]

ESP program staffers believed that local control was consistent with comprehensive local reform. They simply assumed that there was a locus of control from which LEAs could be directed. For example, when ESP sought out local districts interested in a "comprehensive" reform program, recruitment concentrated only on school superintendents. They were seen as the key to local action, and their state of mind was regarded as critical. The views, comments, and wishes of principals and teachers were not even explored.

To insure their administrative control, ESP staff proposed a "rigorous" system of monitoring, while emphasizing that such a system "must respect and support the local project design."[6] ESP

monitors were to work directly with local central administrators in a "partnership." It remained unclear how federal partners were to respect local control if the project did not conform to federal expectations.

Local Control vs. Federal Ambitions

Earlier federal reform programs had been plagued by a lack of local commitment to federal program plans, and by consequent attempts by federal executives to change local programs. Although ESP was designed to overcome this problem, it didn't. The more zealously federal reformers tried to encourage implementation of ESP, the more they interfered with local autonomy, and the greater was local resentment. As one superintendent remarked, "[The federal ESP staff] were attempting to take complete control of a school district through an ESP program by their own methods without regard to the wishes and practices of the school district."[7]

There was conflict between the federal ESP staff and local practitioners from the outset. The latter were understandably interested in obtaining grants, and to this end wished to please ESP Washington. Some had received planning grants based on brief initial letters of interest. But the proposals they then prepared under these grants were more serious than the letters, and had to fit with the local organization. This made it difficult to achieve the comprehensiveness and coherence ESP wanted. As one superintendent explained, the district "wanted it to be our program, but we could never write anything acceptable to [ESP Washington]." Eventually, after several frustrating encounters with the federal ESP staff, this superintendent

simply called the federal project monitor so that she "could tell us the words she wanted us to say."[8] This, of course, involved at least a nominal loss of autonomy, as it eliminated control of project design by local participants. A project director in another site said of his proposal writing. "My problem was that the more they demanded . . . [comprehensive] change, the further it took ownership away from those who developed the project; and the more they asked me to rewrite, the more isolated I became from the staff, because I was really bastardizing their work."[9]

Here and elsewhere, principals and teachers began to resent local administrators, because of what seemed a loss of autonomy. One principal explained, "Like most teachers, I kind of felt that it was a program being pushed on them."[10] And federal pressure often placed local administrators in a position in which they had to make their projects appear to be something they are not. As one project director put it: "Binswanger had a way of telling us what we would want. . . . [His] thing was to get the district to 'do what they wanted,' yet to get the money we had to lie," by making the district seem more coordinated than it really was."[11]

The problems did not diminish as the program took shape. As one local project director said, "We thought it was going to be no government interference and we could do what we wanted with the money. Instead there was constant harassment and interference."[12] One superintendent was initially enthusiastic about what he described as the government's notion of "non-interference" in development of his program. But after he went to Washington--as he thought, to sign

the agreement for funds--he reported: "[ESP Washington] tore our budget apart while we sat there completely unprepared. They threw out this and that and tried to tell us what we needed to do for the money."[13] But what Washington told its local partners often did not fit local plans for reform. Federal insistence on community involvement, for example, meant more demands on and less control for local central administrators. As one superintendent explained, "One of the typical problems with the program was that the federal government wanted the community involved, but the community represents such a diverse population that its involvement only created different factions of interest who wanted different things to happen in school."[14]

In one site, in fact, federal officials saw the failure of local administrators to involve the community in proposal writing as a prime problem. The superintendent explained, "NIE kept pushing the community school idea." He felt that they had already involved the community too much, whereas "NIE kept pushing for more."[15] One federal ESP staff member, after a day of argument with local administrators over their initial proposal, wrote his own proposal outline. Local administrators followed that outline in the next proposal they submitted, and it was approved shortly thereafter.[16] But what "their" proposal now promised was no longer what they had intended. In another case, the federal ESP staff wanted every school in the district to be included in the program, but, as the superintendent explained, there were great difficulties "in trying to get different schools together, because every school wanted to do their own thing." Federal insistence on local involvement reduced local administrators' ability to initiate

change where they thought it might succeed,[17] and reduced central control even more.

Local administrators were often bewildered by federal efforts to shape their projects. According to one superintendent, the ESP staff "cut some components" of local projects "because they were not in some way, a way I cannot define, consistent with what NIE wanted."[18] And a local project director reported: "We would ask [ESP Washington] what to do and they would say, 'you determine it.' And then after we would, they would say it was not right and not give us the money. It was really a Catch-22."[19]

Local admininstrators felt that their autonomy was being compromised. They were being told to develop their own programs at the same time as federal reformers pressed them to meet federal notions about how the change should occur. Washington consistently demanded a coordinated program that ran against the grain of the decentralized local organization. And the tension between federal and local staff was exacerbated because federal administrators seemed unaware of this divergence between federal and local practice.

These conflicts arose mainly from the federal assumption that change could be centrally managed in localities. Yet even when the problems of this belief became apparent, federal reformers never lost their conviction that the local central administration was the key to program coordination. This gave rise to further problems, for federal ESP staff thought it followed that, if control was not being adequately exercised, the local administration was responsible. Thus if reform did not live up to

expectations, the logical next step was for federal practitioners to tinker with the local administration. This they did. One superintendent related how his appointment of a project director, who later became a HEW division director, was vetoed by the federal program staff. He was told that his choice was rejected because "she had conceptualized the program and was not able to implement it in an objective fashion."[20] According to the superintendent in another site, the federal monitor tried to secure the resignation of the project director, to whom NIE attributed difficulties with the program, by threatening to withdraw funds.[21] And the federal staff tried to influence local administrative organization as well as staff. One superintendent reported that the federal program director had ordered what the superintendent described as "reorganizing the whole administration." His response was, "I was the [superintendent] and I was not about to change our organization." Not understanding the reason for this intervention, he saw it simply as a direct contradiction of promises of non-interference by the federal program director.[22] It certainly seemed an odd way to promote local control.

Local administrators often did not share federal monitors' views of the program, and they did not believe that federal objections could be met through federal intervention in the school district. Nearly all local administrators saw their authority

being undermined. As one superintendent explained, after he had been asked to remove the original project director: "On the bright side, it had the effect of unifying the principal and the staff of [the district] because they felt they had been done in by NIE,"[23] but it did not enhance local commitment to the ESP project.

Federal actions thus belied ESP assertions about commitment to local control. Some local administrators gradually abandoned their attempts to maintain control of the program. Where federal monitors were active, local control tended to erode, and with it local commitment. Federal practitioners, for their part, would not or could not accept the decentralized organization of local school districts. They saw the lack of coordination as the result of weakness of will, or intelligence, or both. The federal ESP staff saw local control as a desirable reform of past federal practice, because it would promote local commitment to comprehensive change. But control at the local level is decentralized. While the Washington ESP staff believed that change could occur at a superintendent's direction, local teachers and principals exercise much control independent of the central administration. Managing a local district consists in large part in accommodating these diverse local interests. As a result, ESP, like other federal education programs, took a variety of local shapes. But the Washington staff saw this as a failure of local leadership, while local administrators saw it as an essential feature of their work.

Late in the life of ESP, one senior federal ESP staff member reported that she still saw nothing wrong with the notions of comprehensive change and local control--the problem was in local implementation. What ESP

had needed, she said, was more sincere, dedicated, and capable local school staffers.[24]

Footnotes to Chapter III

[1]This may seem a necessary belief for federal practitioners. Tolstoy, in War and Peace, discusses Napoleon's "command" at the battle of Borodino: "And it was not Napoleon who directed the course of the battle for none of his orders were executed and during the battle he did not know what was going on before him. So the way in which these people killed one another was not decided by Napoleon but occurred independently of him in accord with the will of thousands of people who took part in the common action. It only seemed to Napoleon that all took place by his will." (Tolstoy, L. War and Peace. New York: Washington Square Press, 1963, p. 399.) Administrators, like generals, may command more than they control.

[2]Timpane, M. Educational experimentation in national policy. Harvard Educational Review, November 1970, 40(4), pp. 547-566.

[3]Interview with federal ESP director of evaluation, July 29, 1977.

[4]Interview with former director of the National Institute of Education, October 24, 1976.

[5]Interview with federal ESP staff member, October 29, 1976.

[6]Internal federal ESP program document prepared for the National Institute of Education, dated September 18, 1972.

[7]Interview with former district superintendent of the Edgewood School District, San Antonio, Texas, April 13, 1977.

[8]Interview with district superintendent in Craig City, Alaska, May 4, 1977.

[9]Interview with local ESP project director in the Constantine School District, Constantine, Michigan, May 2, 1977.

[10]Interview with school principal in the South Umqua School District, South Umqua, Oregon, May 24, 1977.

[11]Interview with local ESP project director in the Berkeley Unified School District, Berkeley, California, May 10, 1977.

[12]Interview with local ESP project director in the Carbon County School District, Saratoga, Wyoming, April 21, 1977.

[13]Interview with district superintendent of the Carbon County School District, Saratoga, Wyoming, April 21, 1977.

[14]Interview with district superintendent of the Berkeley Unified School District, Berkeley, California, May 11, 1977.

[15]Interview with district superintendent of the Franklin Pierce School District, Tacoma, Washington, March 30, 1977.

[16]Ibid.

[17]Interview with district superintendent of Supervisory Union #58, Groveton, New Hampshire, June 7, 1977.

[18]Interview with district superintendent of the Franklin Pierce School District, Tacoma, Washington, March 30, 1977.

[19]Interview with local ESP project director in the Edgewood Independent School District, San Antonio, Texas, April 13, 1977.

[20]Interview with district superintendent of the Edgewood Independent School District, San Antonio, Texas, April 13, 1977.

[21]Interview with district superintendent of the Franklin Pierce School District, Tacoma, Washington, March 30, 1977.

[22]Interview with district superintendent of the Carbon County School District, Saratoga, Wyoming, April 21, 1977.

[23]Interview with district superintendent of the Franklin Pierce School District, Tacoma, Washington, March 30, 1977.

[24]Interview with senior federal ESP staff member, November 17, 1976.

IV. KNOWLEDGE AND PRACTICE

Federal school reformers believed that change could be greatly facilitated by new knowledge. They considered the everyday knowledge and skill of local practice to be deficient, and thought that scientific knowledge could help to improve practice. The notion that social science would be useful to local practice and federal policy makers was enthusiastically accepted in the late 1950s and 1960s.

But efforts to improve federal programs and local practice through social science consistently fell short, both in providing information helpful to management and in determining the success of programs. Thus, the 1960s were marked by repeated efforts to improve the application of social science to education. The ESP was one such improvement. It was cited in HEW Secretary Robert Finch's opening comments to the House Committee on Education as one of several programs that were to address earlier scientific shortcomings.[1] One year later, President Nixon, in his message to Congress on "Reform and Renewal" in education, described the ESP as the "missing bridge between educational research and actual practice." The ESP was seen as the prototype of research programs to be contained in the newly proposed NIE.[2]

The ESP staff saw their mission as part of NIE's proposed response to the "general criticism" of educational reform that had been "extended to include previous research . . . ."[3] The ESP staff reasoned that past evaluations "had failed to capitalize on the opportunity to contribute to the knowledge base" because they used traditional social science methods: they typically looked only at outcomes, and found "no significant differences." Being ignorant about what had actually occurred in school programs, these evaluations offered little useful information on the change efforts themselves.[4]

The ESP staff argued that the deficiencies in program evaluation could be rectified by a different and stronger commitment to social science. This was apparent first in its apportionment of 25% to 30% of the program funds for evaluation.[5] Second, this evaluation, unlike others, was not to be conducted on a "fly-in fly-out basis."[6] It promised to use and develop "new and improved research and evaluation strategies" that would be closely attuned to the process of change in schools--how it worked, why it went well or badly, and what helped or hurt. This in turn would allow: "(1) significant improvement in the knowledge base concerning formal education; (2) better understanding of the processes of change in education; (3) more satisfactory investigation into the effectiveness of schooling; (4) the generation of programmatic information that will be of immediate use to practitioners."[7] There was a pervasive belief that local practice could be refined by supplanting or supplementing practitioners' knowledge and skills with social science, and that this could be achieved by using scientific techniques that were "innovative" and "experimental in conceptualization, approach and practice."[8] How the evaluation was to be conducted was less clear; but that social science was the appropriate tool with which to reform local practice was aggressively asserted.

Two characteristics of the evaluation showed the direction of intended reforms. "First, the program required a research vehicle which would maximize what one could learn from a collection of field experiments--summative evaluation and documentation (that is, knowledge about the degree to which 'comprehensive change' was occurring in any school district); and second, the program required the support

of an evaluation designed to assist in the improvement of programs and projects while they were in the process of development and operation--formative evaluation."[9] The evaluation of ESP was to have two related reform criteria: the innovative methods to be used and the new type of information to be produced; and improved management of federal programs at the federal and local levels through the use of program-related information supplied by ESP's evaluation.

To that end, an elaborate three-tiered evaluation scheme was proposed. The major evaluation component (termed "on-site" or "level II" project evaluation[10]) was to be conducted by private firms of professional social scientists. It was to account for two thirds of the monies spent on the evaluation,[11] and to address the main federal concern: to improve local practice by producing a new type of social science knowledge on the process of local change, which would help improve program management.

The ESP staff was particularly concerned with past social science efforts to determine the criteria and judgments about the outcomes of schooling. "Many studies," an early ESP staff memo stated, "can be criticized on the ground that they have equated measurement of a relatively narrow band of cognitive skills with the outcomes of schooling." To address this concern, the ESP evaluation sought "alternative strategies for research" that would "not only have to develop broad-based measures," but would also "have to study particular phenomena over sufficient time and in sufficient depth to allow for more costly methods of behavioral observation as well as paper and pencil tests." The ESP program staff was also concerned that "variables which are traditionally studied have not been adequate

in explaining differences" in local school change efforts. It was therefore proposed that ESP develop a research strategy having "a different and probably more qualitative set of variables to measure differences between schools." By so doing, the ESP staff thought they could figure out what made change work, and thus how to improve local practice.[12]

The information to be produced by the level II evaluation was therefore to be useful to federal and local decision makers. The ESP staff held that if the evaluation, which sought to serve decision makers, was accurate, then it would ipso facto be useful to decision makers. But this assumed that social science knowledge would be preferred to the knowledge that was ordinarily used in local decision making. It was simply assumed that social science knowledge was superior to ordinary or professional knowledge, and thus would be used in decision making.

The second way in which social science knowledge was to improve local program management was through "level I" or "local project" evaluation, "performed by or under the direction of local project personnel" and complementing the level II effort. It constituted a small part of local site budgets and was to be "formative"--i.e., "oriented primarily toward operational decision making and planning."[13] Level I evaluation was to satisfy local project needs "to improve the ongoing operations of a particular project."[14] And it was to minimize the political controversies that had arisen over evaluations in the past. An ESP memo explained, "There is a natural conflict between the needs of a long-term study and the needs of project participants for immediate help, therefore if one wishes to insure both purposes are

served, there are advantages to separating them."[15] It was also thought that locally based evaluation would "seem to be more effective and to have more local credibility."[16]

The third evaluation level was the least developed but most ambitious. Level III was intended to aid, pull together, and summarize the other two levels. It was intended to look "across projects . . . to provide technical/professional assistance and cross-fertilization, to provide informed criticism of the methods and conclusions from each level [of the evaluation], to derive and implement evaluation approaches and models that may be used to assess the merit of comprehensive experimentation, and to identify and formulate evaluation problems which may be turned over to basic research programs for investigation."[17] ESP Washington thought that the major "users" of such an evaluation would be federal policy makers, who would be able to replicate elsewhere the successful programs and evaluation methods identified by it.

Thus, all three evaluation levels rested on certain assumptions about the information each level would provide and the ways in which it would actually assist practice. There was no clarity about how the evaluation was to be implemented, but two central notions were plain enough: it was to produce new knowledge on the local change process; and, through application of that new knowledge in federal and local practice, evaluation would facilitate local reform.

Yet the federal assumption that comprehensive change could be evaluated presented an immediate obstacle to plans for the evaluation.

Since the notion of comprehensive change was itself very unclear, those asked to assess it did not know what to do. This provoked arguments over the definition of comprehensiveness and how to evaluate it, frustration at the enormity of the evaluation task itself, and instability in the evaluation staff. In the end, this innovative evaluation produced rather weak results. It gradually became more traditional, its methods more familiar, and its aims less ambitious.

Struggles with Comprehensive Change and Holistic Evaluation

To improve upon the limited success of past social science evaluations, the ESP evaluation sought to understand better "how the process of change takes place," and to provide practitioners with knowledge about program outcomes.[18] One of its major components was a novel anthropological/ethnographic approach that would allow on-site researchers to study and report on the local change process as it occurred. This documentation of the "natural history" of the project was to deal with

> . . . the project as a totality. The interrelationship among the components of the project and the nature of the forcing [sic] functions, which promote or inhibit successful project development, are to be identified and assessed with respect to relative influence or impact.[19]

It was hoped that such a method of study would provide social science knowledge more useful than that generated by the quantitative focus of previous outcome evaluations, because it would include all aspects of local change and not be narrowly circumscribed by quantitative methods. The expected "holistic" information about the local change process could then be used by federal ESP program managers, as well as by others who wanted to reform local practice.

While a qualitative approach was emphasized, statistical studies also were required. The federal plan was "to include but not be limited to student achievement, attitudes, staff performance and other more traditional categories of evaluation."[20] But as the federal director of the evaluation pointed out, the aim was "to learn how the process of change took place"; he was personally convinced that there was little in conventional quantitative evaluation that was worth doing. This more traditional criterion in the Request for Proposal (RFP) was apparently due to pressure from the HEW office of the Assistant Secretary for Planning and Evaluation, which, in the view of ESP staffers, "did not think an evaluation without pre- and post-testing data was an evaluation."[21]

The burden of devising the non-traditional approach to evaluation fell on the shoulders of the level II contractors. The RFP articulated this assumption: ". . . in the course of the ESP, many needs will be identified for the development of approaches to evaluating educational settings. . . . As the new approaches and tools are developed, it is expected that they will be tried in the field by the evaluation and documentation contractors."[22]

Once federal ambitions for ESP evaluation were translated into an RFP, potential level II contractors came forward with a hopeful and accepting reaction. They expanded the already grandiose federal plans for ESP. Responses typically expressed awareness of the failure of past evaluations, and made vague promises to use social science methods to produce more fruitful knowledge on the local change process. One level II document phrased this in the following

way: "Evaluative researchers have relied primarily on mathematical models (the traditional research models of sociology and psychology) to judge the success or failure of educational programs. . . . This [mathematical] approach to evaluation, which is too narrow in focus and has limited evaluators' ability to do more than speculate about program results, led to the general state of dissatisfaction of those being evaluated. . . ." In its place, this contractor promised a more qualitatively oriented evaluation.[23] One proposal even offered a line of inquiry so novel that it would help create a new profession: "In recent years, evaluation has become a technical specialty with associated methodologies, canons of procedure, fairly abstract bodies of knowledge and theory and accumulated experience, a set of more or less renowned practitioners and a vested interest in developing all of these further. Our own evaluation efforts are expected to make a contribution to this rapidly growing field which threatens or promises to provide grounds for a new and distinct profession."[24]

The evaluation, then, was to build social science knowledge on educational change. But in the absence of a usable definition of that phenomenon, level II contractors quickly came to see the ESP program aim--comprehensive change--as an impossible target for evaluation. As one level II contractor explained, "You must recognize from the outset that this situation made it difficult to do research in a way which would qualify as research to the outside."[25] Another contractor, explaining the low level of productivity in his evaluation, said that the notion of comprehensive change was only "a metaphor" with "no reality" in the world of social science research.[26] A former member of a level

II contractor's staff explained that the original RFP was "awfully vague" and he "did not know how much time [his firm] spent debating comprehensive change!"[27] Another member of the same firm added that his firm later "pushed NIE to define it."[28] A staff member of another contractor referred to the difficulty by saying that "the lack of a defined treatment really came home to roost because it was difficult to know what to measure in terms of outcomes."[29] And to make matters worse, many level II evaluators felt unable to share their doubt about comprehensive change with their client.[30]

The same problem arose in the "ethnographic" component of the level II evaluation. One level II staff member called most of his own work "essentially meaningless in the context of the program."[31] He explained that what ethnographic reporting was done was hard to relate to documenting "comprehensive change," since that notion was so ill-defined. The director of a level II evaluation touched on the problems of communication with federal program people over this issue when he said, "We talked a lot with NIE about the trouble of trying to use NIE's five facets[32] and that we would have preference for social science categories, but that was another area of negotiation."[33] This comment went to the root of the problem: the contractor found the federal ESP notions problematic, but the federal staff did not.

One resulting problem was a lack of federal criteria and coherent guidance for ESP evaluation. Federal managers wanted the evaluation to provide a "holistic" view of local change efforts, yet they didn't know what this meant. They knew only what they didn't want, and seemed to think that by naming an alternative they had created one.

But to call evaluation holistic provides neither criteria for judging what that is or is not, nor guidance for how to do research. Given this lack of guidance, and given that not all aspects of local change efforts could be studied, level II contractors had a problem: which aspects to evaluate, and how? There were no consistent answers for any one contractor, nor consistent answers among them. Indeed, the problem was compounded by the federal ESP staff's interest in exploring various features of comprehensive change. These features changed from time to time and from monitor to monitor, and this created more uncertainty for the level II staffs. Level II contractors were continually trying to do too much at once, or too little, or the wrong thing, or some combination of these. The evaluation was never given a central focus, and as a consequence, work floundered.

Numerous conflicts and debates arose, over what aspects of local change efforts might be appropriate for study, and with what social science tools. But once level II contractors decided these matters in some way, they placed limits on the work, and then federal ESP staff could see evaluation efforts that were incomplete and too specific--for there were always more aspects of local change to be studied. Level II contractors, with the limited tools of social science, and limited time, simply could not provide the "holism"--that fantasized complete picture of local change--that the federal staff dreamed of.

The ensuing demands placed upon level II contractors contributed to instability, confusion, and even capriciousness in the evaluation. The rate of turnover of project directors was high: six of the seven

original directors resigned during the course of ESP. Turnover rates for other level II staff were also high.[34] The lack of criteria for evaluation also allowed federal monitors to change their theoretical and methodological priorities continually. As one level II staff member explained, "ESP had a history of flip-flopping 180°, and as the project officers changed, so did the direction given to the evaluations."[35] Another contractor mentioned the same problem: "Each federal project monitor had their [sic] own idea of what the rural school evaluation should look like. Consequently, they kept asking for changes that would suit their impressions. But often their changes had not been agreed upon earlier . . . and since they could not easily be defined to third party researchers, we stuck to our design."[36] Other contractors, not so self-assured, often changed their design and data collection procedures in response to the demands of a federal monitor.

This situation made it difficult for contractors to formulate evaluation plans that were acceptable to federal reformers. Contractors often had to rewrite evaluation proposals--in one case, a proposal had to be rewritten four times in one and a half years; it was never approved and the contract was eventually put up for rebidding.[37] The difficulties continued throughout the early part of the project. Several level II contractors were never able to produce an acceptable plan, even years after the evaluation was under way. In fact, two of the seven evaluation contracts were put up for rebidding.

In these circumstances there was constant friction between federal ESP staff and level II contractors. The ESP staff saw a great discrepancy

between what they wanted and what contractors produced, and they saw contractors as incompetent, mercenary, or both. One monitor explained rather simplistically that the contractors were motivated by greed and had little "core" interest in the possible fruits of such research.[38] Another common view was that ESP had given "the wrong people contracts," people who "could not deal with the freedom built into the design."[39] While the latter comment has some basis, it cannot explain the often disastrously poor quality of the evaluation reports, the inconsistent federal management of the evaluation, or the connection between the two.

These problems did nothing to improve the evaluation. In several cases, level II contractors, facing the difficulties of trying to formulate an acceptable approach to the evaluation, decided that comprehensive change could not be evaluated. They gradually abandoned the non-traditional and qualitative for the more familiar quantitative social science methods. As one level II staffer explained, "NIE seemed so muddled about certain things, we felt the only thing we could do was try and publish so that the social science community would know what we were doing."[40] Another said it more simply: he and his colleagues just "did their own thing and nobody knew what the right thing was."[41] Given training and professional values, "their own thing" often turned out to be traditional social science.

Thus in many level II evaluation reports, a major element of the federal evaluation plan--documentation of the change process through ethnographic/anthropological methods--yielded to the infamous traditional quantitative approach. In one level II urban evaluation, studies and reports were based on familiar social science categories: "Student Achievement," "Attitude Instrument Study," "Language Dominance Assessment," "Socio-Cultural Bias of California Test of Basic

Skills," "Self-Concept Pilot Study."[42] Another urban site contractor submitted reports on other familiar subjects--"Survey of Parents in the Jefferson School District," a "Study of Curricular Decision-Making in the Jefferson Experimental Schools Program," "The Status of Individualized Instruction Strategies on the Elementary Level K-6 in the Jefferson School District," and "The Effects of Transiency and Social Integration Upon Academic Achievement."[43]

In ESP's closing years, this shift in emphasis from non-traditional and qualitative to more conventional and quantitative methods found a parallel at the federal level. Federal ESP staff became aware that they were not receiving the novel information and insights expected from the evaluation, that in fact they were receiving a very small return on their investment altogether. With the prospect of having little to show for an innovative and costly evaluation, they tended to emphasize a more traditional quantitative approach, which at least might yield some credible results.[44]

This search for credibility at the federal level was intensified by the transfer of ESP, from the Office of Education to the newly established and research-oriented NIE, one and a half years after ESP's start.[45] One program officer explained the transfer as a declaration of "open season on ESP," adding that nobody in ESP, though, really knew how "to play the research game." She was referring to the ESP staff's lack of social science knowledge required to defend the evaluation. She also argued that "ESP was like boiling an egg--you cannot keep pulling it out of the water every five seconds to see

how it is doing." Her views were to no avail.[46] Another federal staff member also noted that pressure was growing: "People . . . [at NIE] felt that the only way you could justify anything is from hard data or . . . they would not think it was a responsible evaluation."[47] Almost by accident, then, the transfer created a point of convergence for level II contractors and ESP Washington staffers that had not previously existed. In one case, a level II contractor suggested, and had approved by ESP Washington, a major reorganization of the study three years after its conception, which cut back on the ethnographic research and offered a more rigorous traditional study of pupil change (in ESP and non-ESP districts).[48] This new orientation toward a quantitative rather than a qualitative study was a substantial departure from the initial intentions of the federal ESP staff, but it did respond to new Washington ideas about evaluation.

Despite federal efforts to improve the evaluation, its products remained undistinguished. The ethnographic work was typically descriptive and had little analytical content.[49] For the most part, the level II staff tried to record and "document" what occurred in the different parts of a school system. The reports seemed to focus arbitrarily on certain events or aspects of a district, describing activities in various schools in a somewhat journalistic fashion. This was not the disciplined analysis and observation expected in research; neither was it the focused picture of events found in good reporting. It was, instead, often vague, diffuse, and undisciplined by any analytic frame. It offered few insights into comprehensive or any other local change processes.

The later and more traditional quantitative work suffered from other limitations. Understandably, it lacked a definition of the experimental school "treatment" studied; the treatment was unspecified, or only vaguely indicated, in most level II reports. In one instance, an evaluation employed a series of comparison studies of ESP treatment and non-ESP control students, yet there was no discussion of how, or on what basis, the treatment and control groups were determined.[50] Another contractor was simply unable to define a treatment, and left the issue unclear in proposals and reports.[51] In yet another case, a seemingly measurable treatment was chosen, but it had no relationship to the federal program notion: the level II contractor, after various efforts to define a feasible ES treatment, decided on the definition of "dollars delivered to schools."[52] Other reports explained simply that the ESP local project staff "designated certain grades" as experiencing ES treatment;[53] or that ES students represented a sample "not in the sense that they were systematically chosen but that they are a subset of the total number of students who were available for retesting or who were actually tested."[54] In one study, control schools were selected on the basis of geographic proximity to ES treatment schools and willingness to participate in the study.[55]

Even when student populations were identified (by whatever unscientific means), they characteristically became unstable, or incomplete, in what they were to represent. In one study, 263 students were tested out of an initial population of 862 "ES students." The report also implied that there was no assurance that these 263

students were part of the initial ES population, owing to poor records on student enrollment, absenteeism, turnover, and dropouts.[56] Another report explained that the total number of students tested in the treatment and control groups ranged from 50% to 85% of their respective total populations identified each year over a four-year period. There was no indication that the students in either group were part of the original first-year sample.[57]

Some level II contractors simply did not produce reports, in part because of their own awareness of such severe shortcomings as those cited above. Rather than produce work of poor quality, they chose to submit no reports, or purely theoretical reports on social science research issues only vaguely related to ESP. A memo from an NIE evaluation monitor indicates this tendency in a level II contractor's work two years into the evaluation. It reads: "As was feared . . . the conceptual model for organizational studies and analysis continued to be very abstract and failed to deal with the concrete operationalization of the research dimension. At this point the model itself as well as . . . [the] ability to implement it remains unknown."[58] The work of the level II contractors was summed up by a potential contractor bidding for the level III evaluation. The proposal read:

> No mechanism presently operates for returning to the research community or policy makers anything more than an expensive set of case studies on local history.[59]

In summary, the attempt to produce new knowledge on local change was a fiasco. One reason for this was naivete about social science methods on the part of the federal ESP staff. Another was the problem of evaluating comprehensive change, or of carrying out a "holistic" evaluation. In addition, there was weak federal management of and guidance for the evaluation, due in good measure to the combination of grandiose and naive thinking that marked the program's management. ESP's major promise--that change can be comprehensive, and that it can be evaluated holistically with social science methods--remained unfullfilled. The evaluation reports produced by level II contractors were, with a few exceptions, completely undistinguished, and many were an embarrassment.

Management and Knowledge

This social science knowledge also was supposed to improve federal and local management. The ESP staff reasoned that past evaluations had failed in this regard partly because the social science knowledge had been of the wrong sort, and partly because evaluations gave rise to political conflicts between federal and local practitioners. They proposed two remedies. First, the evaluation was to produce knowledge relevant to the needs of practitioners--to "provide the practitioner with adequate information to allow him to judge whether or not changes in his practice are warranted."[60] And second, to reduce conflict and suspicion about the nature of the social science knowledge produced, evaluation was to

take place at three levels, presumably geared to separate local and federal program management needs.

The attempt to design an evaluation that would facilitate program management foundered, however, on differences between federal ideas and local practice. Federal reformers believed that the new knowledge emerging from the different evaluation levels could be helpful to federal and local practice, but local practitioners did not respond as expected. They had knowledge, ordinary and professional, that seemed adequate, and saw little use for the research. Moreover, they saw federal efforts to generate social science knowledge from the evaluation as an instrument of political oversight for their federal managers. This created tension between federal and local practitioners and set the stage for much conflict. Thus, the evaluation affected management of the program, but it hindered rather than helped.

The Level II Evaluation

Local administrators not only showed little interest in or use for the social science findings produced by level II contractors; typically, they saw the level II evaluation as an independent research project unrelated to local operations. Often they regarded it as a nuisance, since the evaluation required much local time and cooperation. Thus, in sites where the social science findings of level II contractors were made available, there is little evidence that they were helpful to local management concerns and interests. One administrator explained that he looked at the level II reports "belatedly" because "they got into gossip rather than the real issue." He added that he had his own inside information on what was going on in the

school districts and did not need to be informed about the "popular image" of how things were.[61] Social science knowledge did not correspond to what local participants believed was going on.

The low utility of the level II evaluation for local administrators was reflected in lackadaisical attitudes toward facilitating data collection. One superintendent, who kept getting letters from the level II contractor "asking us to help get the questionnaires filled,"[62] said that he responded only out of deference to Washington. Local administrators had their own ways of gathering knowledge through local staff, and they saw little to be gained from impersonal questionnaires. Furthermore, their own decision making depended upon considerations and factors that were peculiar to their particular situation, and seemed worlds apart from more universalistic social science methods of gaining knowledge.

Teachers had a similar reaction. They saw level II requests for information as irrelevant, of little use to classroom management. One elementary teacher complained about questionnaires asking, for example, how much time was spent on discipline in class, or how many students in the class smoked. She couldn't see how the answers would in any way help her in the classroom. As a result, she filled out the first one but "otherwise I threw them in the fireplace."[63] Another teacher said, "I don't know much about the evaluation; it bores me."[64] Teachers rarely read the reports issuing from the evaluations, and when they did, the response was the same. One teacher noted, "I saw them but I did not understand them and I

did not know how to read them."[65] The knowledge and skill required for teaching practice overlapped little with the social science knowledge of the evaluations, nor did the second illuminate the first.

Local administrators thus often found it difficult to facilitate level II data collection, even when they wished to. Teachers, principals, and students were already subject to testing and other research efforts, either from other evaluations or from local testing programs, or both. Many resented and resisted the ESP program. Their poor response to level II requests was understandable, but it impeded the work of local administrators, who often were caught between the demands of Washington and level II contractors for data, and their own staff's hostility or indifference to another evaluation.

But if most local practitioners had little use for social science knowledge, the evaluation was not irrelevant. Many of them saw the level II evaluation as a way for ESP Washington to pass judgment on and manage local program activities. This produced considerable tension between federal and local administrators, and contributed to more than a few management problems. Locals saw level II contractors as federal agents.[66] One local project director said that the level II evaluation was part of the government's "experiment in schools," and felt "we were guinea pigs."[67] A local ESP staff member said it was his "impression that the level II people were almost like a fiscal agency representing Washington at the site."[68] A level II project director explained that the local sites he was evaluating had difficulty even telling "the difference between NIE and the CIA,

never mind the difference between my own organization and the federal government."[69] This view of the level II evaluator as a federal agent was pointedly illustrated in a rural site whose community and school staff objected to government-sponsored evaluation because they suspected that the level II questionnaires on the morals and attitudes of their children would eventually be used by the government for measuring antisocial behavior--"another version of Orwell's 1984."[70] Local practitioners typically saw the level II evaluation in adversarial terms, as serving federal purposes.

As local sites were often unsure of what ESP Washington wanted their programs to be, they were particularly concerned about the effect of level II contractors' reports on their funding and their relationship with Washington. As one local person described her site's reaction, they "feared the evaluation would ultimately limit what they could do."[71] The director in an urban site explained that, particularly at the beginning of the project, there was a great deal of hostility toward contractors who "apparently had not even read our proposal before they submitted their plan for evaluation--it was hard to understand how they could make judgments about the project."[72] Local practitioners typically felt that the level II evaluation reports were out of kilter with their own knowledge of the program, and criticized contractors' interpretation of program objectives and implementation.

This divergence in views fed local suspicions about the evaluation, produced extended arguments over the approach and findings of contractors, and fueled political conflicts. In one project, the

local staff and community went so far as to demand the right to review tests before they were given, because they felt that the project's real goals were not being measured.[73] In another site, the project director said that the level II evaluators never quite understood what was going on in the district, and that their reports "were so subjective that they only made me angry--they reported things that never happened and took things that were entirely out of context."[74] In a third site, anti-establishment sentiment in the community coincided with the program staff's concern over a "government-sponsored" evaluation and virtually precluded agreement on what, if any, evaluations should take place. The level II director reported that this attitude "immobilized any research."[75]

The federal ESP staff were in a most difficult situation, for they were hard pressed to judge the validity of local criticisms. While they were skeptical about both the evaluators and the conduct of local projects, they could rarely spend enough time on site to form their own conclusions. In fact, they were busier trying to resolve conflicts between level II contractors and local school people than absorbing contractors' information on the programs. Typically, they tried to walk a narrow line between local complaints and criticisms, and level II responses and counter-accusations. The evaluation component of ESP thus hardly contributed constructively to federal program management. According to the federal design, social science knowledge gained from the evaluation would aid in program decision making and management, but the federal ESP staff certainly never relied on the evaluation, or other social science knowledge, in their own

program management. Indeed, they questioned the credibility of the level II contractors' social science knowledge, which was often inconsistent with what they knew, and relied instead on their own professional and ordinary knowledge in managing ESP. One federal staff member explained that "the evaluation unnecessarily complicated" the program effort and "just did not make sense or a dent in the day-to-day operations."[76]

Ironically, the conflict also made level II contractors reluctant to supply Washington with what findings they did have on local programs, for they feared that federal managers would use such information in their management of local projects, and that this would further strain the evaluators' relations with local staff. The contractors recognized that a workable relationship with their sites was essential, and so argued that their major purpose was to conduct research "in the school site which would take place over a five-year period, . . . not to give NIE information about the project."[77] The reasons for this view were made explicit by one level II director; "We became very leery of our ability of keep our commitment to the sites if we gave much information to Washington."[78]

The reaction of ESP Washington to the situation was deepening concern about whether the evaluation work was in fact being done. In some cases it was not, as a result of difficulties mentioned earlier that level II contractors did not care to share with ESP Washington. In other cases, work completed by contractors was sent on to ESP Washington only after it had become irrelevant to local events, and was of little value to federal monitors. The level

II contractors were in a no-win situation: if they passed pejorative information on to the federal ESP staff, local sites might prevent further data collection and make it impossible for them to fulfill their contracts; yet if they did not share at least some of the information, the federal staff might suspect that they were not fulfilling their contracts. The decision to withhold information so as to maintain an amicable relationship with local sites thus created a new area of conflict with ESP Washington, and a new management problem. The national staff responded by pressuring contractors to divulge information on local sites.

These problems arose from evaluation, and were not resolved by it. They differed little from those found in other evaluations. They are perhaps best illustrated by the events surrounding the federal desire to review level II ethnographic studies, especially those of rural sites. The ethnographers, who were suspect in local sites to begin with, were particularly concerned that if information they gathered was seen by ESP Washington, it might be used against local sites, and that their relations with the communities would be jeopardized.[79] Consequently, individual ethnographers even hesitated to give their information to their own (level II) employers, for fear it might be passed on to ESP Washington.[80] ESP Washington, on the other hand, wanted to be sure that the work was being done, and was of reasonable quality. They feared that the ethnographers had "gone native" and were holding back information.[81] To resolve this problem, an elaborate review system was eventually worked out: the work of the ethnographers would be reviewed only

by people outside ESP Washington, and the reviews would be an assessment only of quality, rather than of substance.[82] This system added yet another management burden, and cost, to the federal ESP staff. It caused extended delays in ESP Washington's examination of level II contractors' work, and of course rendered the information utterly useless for ESP management.

A parallel conflict arose in connection with the plan that level II evaluation should provide information useful to local sites. The original federal notion was that this would minimize political problems, as local sites would receive information from the same source as the federal ESP office. Instead, level II contractors again hesitated to transmit information to local sites for fear of damaging the relationship, and this increased local resentment of the evaluation. As one superintendent explained, the failure of level II contractors to feed information back to their sites was, in his mind, "the height of foolishness." He found it particularly infuriating because they "burned the territory once or twice a year asking the community questions and then not giving answers [sic] to us!"[83] Another project director reported that he and his staff did not have "the slightest idea of what they are doing. The only thing we got out of [the level II contractor] was the Iowa [Test of Basic Skills] Printout." He added that he was told he would "never see any of the information, so don't ask for it."[84] The reticence of level II contractors did not relieve local administrators' suspicions that the evaluation was inaccurate and misleading.

Some level II contractors dealt with this problem by issuing two

separate sets of reports. One level II project director explained that he sent local schools special reports, containing little information, because he feared that if the site personnel saw the real reports they might cut off further data collection. So in place of the real reports he created and sent an "innocuous report or two to keep them happy."[85] Another level II director did the same, simply giving sites some general material for "a political purpose"--so that they would not contest the information or want a part in the data collection and analysis. From the level II contractors' point of view this was critical, "as a strained relationship to the local sites would completely blow our effort to do a summative evaluation."[86] The crazy politics of the evaluation thus meant that information given to local districts could not be useful to them, for if it was it would be unhelpful to the evaluators. The information passed on to the districts was intentionally made useless, so that the evaluation could proceed. In practice, the idea that social science knowledge would provide a management tool for local districts resulted in the deliberate creation of social science fictions, in order that the relationships between evaluators and locals could be made manageable.

The Level I Evaluation

The level I evaluation shared many of these problems. It was to have been locally operated, to help districts improve project management and reduce political conflict. But local sites, particularly rural ones, typically did not understand the notion of creating a formal system of social science knowledge. As one rural school director explained, he knew "nothing, absolutely nothing" about the level I evaluation except that "ESP in Washington said we had to have it." He

added that "the people in the project had no idea what it was for. They were practitioners. Nobody did research . . . . They were normal people and could not be expected to do a formative evaluation."[87] Similar views were expressed even by urban sites, which were more familiar with evaluation. One project director said, "Quite frankly, I did not know what [level I evaluation] meant, but we tried to make it part of the project."[88]

Other local managers were passive about or resistant to level I evaluations because they saw little value in the knowledge to be provided.[89] Thus, they recruited local people as evaluators who had little experience in the area of evaluation.[90] Federal managers were not pleased, either with such selections or with the work produced, and even tried to help find better-qualified people.[91] They saw the level I evaluation as a way of helping local sites to make program-related management decisions, but few locals saw any use for it. The harder the federal staff pressed the locals to take the evaluation seriously, and the more actively they tried to set it up, the more local sites suspected that it was just another way for the "feds" to evaluate them. The federal staff could not believe that local practitioners had little or no need for social science knowledge, even if it was their very own.

In some sites, no level I evaluation worth mentioning took place. In those local sites where some sort of evaluation took place, it often caused additional problems. Teachers and other local staff were restrained, apathetic, or hostile. One level I evaluator, for example, talked about how difficult it was for local staff to "write their objectives." He

explained that "this was not typically the concern of teachers and educators, that is, to systematically formulate their process, because they are used to operating intuitively and by the seat of their pants." The skill and knowledge that worked in the classroom, he believed, was not that of social science. And there was the further difficulty that, since gathering social science knowledge was described at the federal level as an evaluation, "evaluating will always serve as a threat to teachers."[92] His reaction was not unique. Local teachers were disturbed over the amount of data collection, and level I evaluators found it difficult to obtain information on the program because of the resulting local staff resistance.

As a consequence the reports that did emanate from level I were hardly used for the local management of ESP. Typically, only local administrators were even aware of the reports. As one level I evaluator said, "Never has the evaluation been taken seriously in the central office."[93] Other central administrators did use the evaluation, but generally to confirm what had already been decided or known informally. One superintendent explained that it often "helped to confirm our intuitions and feelings."[94] Most, however, found it a bother.

On the whole, then, level I evaluation was not a great success. In some cases, it seemed so irrelevant that it was never even implemented. In others, it added another area of conflict between federal and local managers, and contributed to local administrators' internal management problems because of local staff resistance. At the core of these problems was the erroneous assumption that local

administrators would use social science knowledge as an aid to management, especially if the evaluation was locally based. Instead, local practitioners had their own sources of knowledge, and had little use for level I evaluation. Typically they saw it merely as another way for federal staff to monitor local efforts.[95]

Conclusion

The ESP evaluation scheme was intended to provide new social science knowledge on the local change process, and to facilitate program management. But this scheme encountered the same problems as did other federal evaluations. The social science knowledge gained from the evaluation was of poor quality, and it contributed little to practice at any level of education. Its only appreciable effect was to increase conflict within the ESP program, and to raise management problems all along the line. Indeed, in this sense it created a new area of practice: managing unusable but potentially problematic social science. This has become a large element in the management of evaluation in other federal programs, including Title I and PL 94-142.

Ironically, though, the problems that beset the ESP evaluation did not stem from the poor quality of the knowledge obtained. We suspect that the opposite may be true: knowledge of higher quality might have intensified rather than diminished political conflict. In the local world of practice, social science knowledge--however good it may be, and however useful for other purposes--will not replace other trusted, ordinary, ways of knowing. But while the social science knowledge produced by evaluation is of little use in program management, practitioners fear it will be used by others to limit their freedom of action. The

more effective the information, then, probably the greater the threat. If the ESP studies had had greater credibility, they would thus have received even more attention. This would have increased local suspicion and resistance, political tensions between outside evaluators and local program operators, and federal defensiveness. Conflict would likely have been exacerbated. This raises questions about the long-standing assumption of federal policy that social science knowledge relevant to a program will be useful to program management.

Footnotes to Chapter IV

[1]Finch, Robert. Statement of the Secretary of Health, Education and Welfare, prepared for the House Committee on Education and Labor, 91st Congress, 1st Session. Congressional Hearings, March 10, 1969.

[2]President Richard M. Nixon's Special Message to Congress on education reform, entitled Reform and Renewal. Delivered to the 91st Congress, 2nd Session, on March 3, 1970. Public Papers of the President of the United States, Document #91-267, pp. 228-238.

[3]Internal federal ESP staff program paper for the National Institute of Education, "Statement of the Problem," undated.

[4]Internal federal ESP staff program paper for the National Institute of Education, "What are the Shortcomings of Traditional Evaluation Approaches?", undated.

[5]This compared, for example, to approximately 5% of Title I monies spent on evaluation over its seven year history. McLaughlin, M. Evaluation and Reform: The Elementary and Secondary Act of 1965, Title I. Cambridge, MA.: Ballinger, 1975.

[6]Interview with federal ESP director of evaluation, July 27, 1977.

[7]"Statement of the Problem," op cit.

[8]Internal federal ESP staff program document prepared for the National Institute of Education, untitled, dated September 18, 1972. The purposes were also paraphrased in a federal ESP staff paper to the National Institute of Education, entitled "What Were the Central Characteristics of the Original Approach to the Evaluation Component of the Program [ESP]?", undated.

[9]Ibid.

[10]This component of the evaluation was to be conducted by independent private contractors. There were seven "level II" contractors in all, one for each of the five urban sites, one for three urban sites and one for the ten rural sites.

[11]Internal federal ESP program document prepared for the National Institute of Education, dated September 18, 1972, op cit.

[12]"Statement of the Problem," op cit.

[13]"What Were the Characteristics of the Original Approach to the Evaluation Component of the Program?", op cit.

[14]Internal federal ESP staff program document prepared for the National Institute of Education, dated September 18, 1972, op cit.

[15]Ibid.

[16]Ibid.

[17]"What Were the Central Characteristics of the Original Approach to the Evaluation Component of the Program?", op cit.

[18]Interview with federal ESP director of evaluation, July 26, 1977.

[19]Original federal ESP program RFP for the level II evaluation, "Evaluation and Documentation of the Experimental Schools Project--The First Three Urban Sites," undated.

[20]Ibid.

[21]Interview with federal ESP director of evaluation, July 26, 1977.

[22]"Evaluation and Documentation of the Experimental Schools Project--The First Three Urban Sites," op cit.

[23]Doyle, Wayne, et al. Final Report of the Northwest Educational Regional Laboratory, The Birth, Nurturance and Transformation of an Educational Reform. Portland, Oregon: Northwest Regional Educational Laboratory, August 1976, p. 5.

[24]Report submitted by a division of Scientific Analysis Corporation to the ESP office in the National Institute of Education. A Conceptual Overview of the Evaluation and Documentation Effort. Berkeley, CA.: Scientific Analysis Corporation, December 20, 1972, pp. 11-12.

[25]Interview with staff member of a level II contractor, Abt Associates, Cambridge, Massachusetts, March 29, 1977.

[26]Interview with director of a level II evaluation, Scientific Analysis Corporation, Berkeley, California, August 31, 1977.

[27]Interview with staff member of a level II contractor, Northwest Educational Regional Laboratory, Portland, Oregon, December 2, 1976.

[28]Interview with staff member of a level II contractor, Northwest Educational Regional Laboratory, Portland, Oregon, April 14, 1977.

[29]Interview with staff member of a level II contractor, Abt Associates, Cambridge, Massachusetts, March 24, 1977.

[30]Interview with director of a level II evaluation, Northwest Educational Regional Laboratory, Portland, Oregon, December 2, 1976.

[31]Interview with staff member of a level II contractor, Northwest Educational Regional Laboratory, Portland, Oregon, December 2, 1976.

[32]See beginning section on the federal definition of comprehensive change. This definition listed five facets (i.e., curriculum, organization, instruction, administration and community) which comprise comprehensive change.

[33]Interview with director of a level II evaluation, Abt Associates, Cambridge, Massachusetts, August 16, 1977.

[34]Interview with staff member of a level II contractor, Northwest Educational Regional Laboratory, Portland, Oregon, December 2, 1976.

[35]Interview with staff member of a level II contractor, Aries Corporation, Minneapolis, Minnesota, June 3, 1977.

[36]Interview with staff member of a level II contractor, Abt Associates, Cambridge, Massachusetts, March 24, 1977.

[37]Telephone interview with director of a level II evaluation, Human Interaction Research Institute, a subcontractor of the Scientific Analysis Corporation, Berkeley, California, August 31, 1977.

[38]Interview with senior federal ESP staff member, November 17, 1976.

[39]Interview with federal ESP director of evaluation, July 27, 1977.

[40]Interview with staff member of a level II contractor, Abt Associates, Cambridge, Massachusetts, March 24, 1977.

[41]Interview with staff member of a level II contractor, Northwest Educational Regional Laboratory, Portland, Oregon, December 2, 1976.

[42]Reports submitted or in preparation for submission to the ESP office in the National Institute of Education by level II contractor, D. C. Development Associates of San Antonio, Texas.

[43]Reports submitted to the ESP office in the National Institute of Education by level II contractor, Northwest Educational Regional Laboratory, Portland, Oregon.

[44]Interview with director of level II evaluation, Scientific Analysis Corporation, Berkeley, California, May 12, 1977; interview with director of level II evaluation, Aries Corporation, Minneapolis, Minnesota, June 3, 1977.

[45]Interview with federal ESP staff member, December 3, 1976.

[46]Interview with senior federal ESP staff member, November 17, 1976.

[47]Notes on a telephone conversation with federal ESP staff member on September 19, 1977.

[48]Interview with director of a level II evaluation, Abt Associates, Cambridge, Massachusetts, August 16, 1977.

[49]This view does not include the portion of ethnographic work to be produced by the rural level II contractors, Abt Associates, because we have not yet seen it.

[50]Report submitted to the ESP office in the National Institute of Education by level II contractor, D. C. Development Associates of San Antonio, Texas.

[51]Based on an on-site review of the work of a level II contractor, Scientific Analysis Corporation, Berkeley, California.

[52]Abt, Wendy. Design Issues in Policy Research: A Controversy. Interim Report #1 on Abt Associates' Pupil Change Study. Cambridge, MA: Abt Associates, February 27, 1976.

[53]Welch, Richard. Academic Achievement of Mexican-American Students. San Antonio, Texas: D. C. Development Associates, April 1977, p. 11.

[54]Moore, Vernon. The Third Report (Vol. 1) (Research Report). Washington, D.C.: The National Urban League. Submitted to the office of the ESP in the National Institute of Education, April 30, 1975.

[55]Interview with staff member of a level II contractor, Abt Associates, Cambridge, Massachusetts, March 24, 1977.

[56]Moore, Vernon. Student Outcome Report (Vol. 1) (Research Report). Washington, D.C.: The National Urban League, Spring 1975, p. 6.

[57]Cervantes, R., & Bernard, H. Psychosocial Growth and Academic Achievement in Mexican-American Students. San Antonio, Texas: D. C. Development Associates. Submitted to the office of the ESP in the National Institute of Education, March 1977.

[58]Internal federal ESP memo from Ray Coward to Norman Gold on the Northwest Educational Regional Laboratory evaluation, dated February 22, 1974.

[59]Proposal submitted by Contract Research Corporation to the National Institute of Education. Level III Evaluation, Documentation and Analysis: Technical Proposal. Cambridge, MA.: Contract Research Corporation, May 28, 1978, Chapter I, p. 4.

[60]Federal ESP program RFP for the level II evaluation and documentation of the Experimental Schools Project, undated.

[61]Interview with local ESP project director in the Edgewood Independent School District, San Antonio, Texas, April 13, 1977.

[62]Interview with district superintendent of the South Umqua School District, South Umqua, Oregon, May 22, 1977.

[63]Interview with two teachers in Supervisory Union #58, Groveton, New Hampshire, June 16, 1977.

[64]Interview with teacher in the Southeast Alternative School District, Minneapolis, Minnesota, May 31, 1977.

[65]Interview with teacher in the Edgewood Independent School District, San Antonio, Texas, April 12, 1977.

[66]The same phenomenon was experienced by the researchers for this project, despite a concentrated effort to make a distinction between the federal government and the researchers' own independent agency. Inevitably, the researcher was introduced as being from NIE or Washington.

[67]Interview with local ESP project director in Supervisory Union #58, Groveton, New Hampshire, March 22, 1977.

[68]Interview with local level I evaluator in the Edgewood Independent School District, San Antonio, Texas, April 11 & 13, 1977.

[69]Interview with director of a level II evaluation, Abt Associates, Cambridge, Massachusetts, August 16, 1977.

[70]Interview with district superintendent, Craig City, Alaska, May 4, 1977.

[71]Interview with teacher in the Southeast Alternative School District, Minneapolis, Minnesota, May 31, 1977.

[72]Interview with local ESP project director in the Southeast Alternative School District, Minneapolis, Minnesota, March 22, 1977.

[73]Interview with staff member of a level II contractor, Aries Corporation, Minneapolis, Minnesota, June 3, 1977.

[74]Interview with local ESP project director in the Edgewood Independent School District, San Antonio, Texas, April 13, 1977.

[75]Interview with director of a level II evaluation, Human Interaction Research Institute, a subcontractor of Scientific Analysis Corporation, Berkeley, California, August 31, 1977.

[76]Interview with federal ESP staff member, October 28, 1976.

[77]Interview with staff member of a level II contractor, Northwest Educational Regional Laboratory, Portland, Oregon, April 14, 1977.

[78]Interview with director of a level II evaluation, Abt Associates, Cambridge, Massachusetts, August 16, 1977.

[79]Interview with ethnographer for level II contractor, Abt Associates, Cambridge, Massachusetts, April 19, 1977. Also see Herriot, Robert. Ethnographic case studies in federally funded multidisciplinary research: Some design and implementation issues. Anthropology and Education Quarterly, May 2, 1977, 8, pp. 106-115.

[80]Interview with ethnographer for level II contractor, Abt Associates, Cambridge, Massachusetts, May 24, 1977.

[81]Interview with director of a level II evaluation, Abt Associates, Cambridge, Massachusetts, August 16, 1977.

[82]Conversation with federal ESP staff member, October 11, 1977.

[83]Interview with district superintendent of the South Umqua School District, South Umqua, Oregon, May 23, 1977.

[84]Interview with local ESP project director in Supervisory Union #58, Groveton, New Hampshire, June 9, 1977.

[85]Interview with director of a level II evaluation, Northwest Educational Regional Laboratory, Portland, Oregon, May 25, 1977.

[86]Interview with director of a level II evaluation, Abt Associates, Cambridge, Massachusetts, August 16, 1977.

[87]Interview with local ESP project director in Supervisory Union #58, Groveton, New Hampshire, March 22, 1977.

[88]Interview with local ESP project director in the Franklin Pierce School District, Tacoma, Washington, March 30, 1977.

[89]Interview with district superintendent of a local rural school site (anonymity requested), May 2, 1977; interview with district superintendent, Craig City, Alaska, May 4, 1977.

[90]Interview with local level I evaluator, Southeast Alternative School District, Minneapolis, Minnesota, May 31, 1977.

[91]Interview with local level I evaluator, Piedmont School Project, Greenville, South Carolina, May 18, 1977.

[92]Interview with assistant district superintendent and director of local level I evaluation in the Franklin Pierce School District, Tacoma, Washington, March 30, 1977.

[93]Interview with director of local level I evaluation in the Edgewood Independent School District, San Antonio, Texas, April 11, 1977.

[94]Interview with district superintendent in the South Umqua School District, South Umqua, Oregon, May 23, 1977.

[95]After the failure of both the level I and level II evaluations, the level III evaluation never was set under way. It was initially designed to rely on the social science findings of the level I and level II evaluations as a base for providing feedback to

federal and local managers. But because of the poor quality of social science findings from the level I and level II evaluations, a level III evaluation hardly seemed possible. Its implementation also promised more of the same as found in level I and level II: unused social science findings and a new layer of conflict and management problems at both the federal and local levels. For example, one local superintendent, when he heard about a level III evaluation, expressed what seemed to be a general local feeling: he wrote to ESP/Washington and suggested that there be a level IV evaluation, because he felt "the hell with them evaluating us all the time; I decided we should evaluate them for a change." (Interview with district superintendent of the Carbon County School District, Saratoga, Wyoming, April 21, 1977.)

From the level II contractors' perspective, a level III study seemed to invite more tensions in their already fragile relationship with federal and local managers. Consequently, when the federal staff raised the possibility of a level III evaluation, they received "a lot of flack" from level II contractors. (Interview with director of a level II evaluation, Abt Associates, Cambridge, Massachusetts, August 16, 1977.)

The general federal feeling toward a level III evaluation was expressed by one federal monitor: "Why throw good money after bad?" (Interview with former federal ESP evaluation director, October 29, 1977.) Given federal opposition to the level III evaluation, at least as initially intended, it was finally dropped with the departure of its one prominent supporter, the original program director. (Interview with federal ESP staff member, October 24, 1976.) This decision came as a welcome relief to level II contractors and to those local sites who had been aware of its possible implementation. (Interview with director of a level II evaluation, Abt Associates, Cambridge, Massachusetts, August 16, 1977; interview with district superintendent of the Carbon County School District, Saratoga, Wyoming, April 27, 1977.)

V. CONCLUSION

The Experimental Schools Program was in many important respects similar to the Great Society school reform programs it sought to improve upon. It aimed to solve the problems of children from poor or minority-group backgrounds. It sought to do this by extensively reforming local practice. The instruments for reform were federal money, ideas, and direction. In other respects, ESP differed from the Great Society school programs. It was small, for one thing. For another, it was able to spend more per pupil than most earlier programs. And finally, its managers were free from many of the political and administrative constraints familiar in other programs, because its funds were discretionary. ESP monies were tied to no legislation that specified how much money was to be spent, or on whom, or for what purposes. These differences should have made it easier for the program to achieve its ambitious goals. They had been deliberate, after all, part of an effort to design a program that would overcome the drawbacks of earlier programs. Yet by our account, and in the view of many of the federal managers, the program failed. It attained none of its goals, including its intention to improve on the record of those earlier federal programs that ESP managers regularly criticized. Why? And what lessons may be learned?

If ESP shared the fate of its predecessors, it was in part because the program rested on naively ambitious notions about school reform. ESP aimed, for example, at "comprehensive change," ignoring the practical wisdom of administration as well as a good deal of social science,

both of which suggest that change in human organizations characteristically occurs in small, often unrelated ways. These many incremental steps sometimes add up to large changes, but they consist of many modest unconnected steps. They are almost impossible to predict, or to plan. The ESP program managers had an extraordinarily overdeveloped faith in local school managers' ability to devise and implement a grand plan for integrated change. In this they resembled their brethren in the Great Society programs, who seemed to believe that ideas for reform are better when they are larger and more systematic. Such ideas are perhaps for that reason more appealing to those whose intellectual formation occurred in the confines of formal education, where the scholastic values of order and tidiness prevail, and where teachers value above all the ability to make synthetic formulations. But managing a complex organization is not analogous to writing a good term paper. The skills typically rewarded in the latter are often an impediment to the former.

It might also have been this curiously scholastic turn of mind that encouraged federal program managers to misunderstand the character of local school organizations. They behaved as though local educational authorities were organized in a centralized and hierarchical fashion, as they might be pictured in a study of the schools' formal organization. Federal managers made demands on school superintendents consistent with such a tight picture of local organization. But local school systems are neither tight nor tidy. The consist of individual schools that are at least semi-autonomous with respect to the central administration, and of teachers within each school who are at least

semi-autonomous with respect to that school's administration. Each local practitioner is free to go his own way, in good part. Any change that begins at the center in such a setting, even if it be modest, is likely to be interpreted in many different ways at the periphery. By pressing local superintendents to devise and carry out comprehensive change, federal managers were trying to produce a sort of orderly behavior quite unfamiliar to local organizations. For even if some LEAs could plan comprehensively, they would rarely be able to act in a coordinated way to implement a broad and integrated plan for change. Such things do happen occasionally, but only when the pressures for action are extraordinary. ESP, however, was not the focus of such extraordinary pressures. It was only a program of discretionary aid, based on freely contracted arrangements between federal and local officials. By behaving as though local school districts actually worked as an organization chart might suggest, federal officials encouraged more resistance to their plans than might have been generated by other approaches.

This synthetic, almost academic turn of mind was also evident in the ESP program managers' efforts to reform the information that local and federal officials used in making decisions. These managers believed that if they devised a new and more comprehensive scheme for social science evaluation, reliable scientific information could replace the knowledge commonly used in managing federal programs and local schools. They ignored the great frailties of social science, and the strengths of both ordinary knowledge and the accumulated professional knowledge of teachers and administrators. Like many other federal reformers of

the last two decades, ESP program managers assumed that social science can provide authoritative guides to action, and can strongly influence what teachers and administrators do. In fact, social science is rarely authoritative in that sense, and particular studies usually influence particular decisions only weakly. The ESP evaluation was costly and elaborate, but not irrelevant. While it produced information that was mostly of poor quality, and not helpful to teachers and administrators, it did add another focus for federal-local conflict.

One lesson from this costly affair is that federal competence in education, whatever else it may be, does not rest on technically superior knowledge. Federal administrators are far removed from local practice, both by geographical distance and by the distances introduced by differences in the work of federal and local managers. Applied social research does nothing to lessen these distances, and in the case of the ESP, research increased the distance. Federal managers were enthusiastic about applied research in part because they had big ideas about local reform, but little fiscal and political leverage on localities. The presumed authority of social science was no substitute for political authority, though, at least in this case. And the fruits of the research were not such as to change any local minds about how education should be carried on. As an instrument of persuasion, social research was a flop in the ESP.

Another lesson is that federal competence in education does not extend, even in easy cases like ESP, to detailed guidance to localities. The program had no legislative constraints, and it engaged only localities that wished to participate. Yet such was the distance between Washington

and local schools that the efforts of national staff to offer any more than general guidance about reform were typically fruitless, and often problematic. Admittedly, one cannot generalize freely from the case of ESP to other programs--the ESP reform ideas were extraordinarily ambitious, even in the lexicon of federal education policy. But it does not seem precipitous to argue that if detailed guidance doesn't work in such a modest, unencumbered, and well-funded endeavor as ESP, it is unlikely to work in larger, less well-staffed and -funded, and much more overburdened programs.

This is not to deny that social science and federal reform ideas have something to offer education. In the case of social science, however, federal officials and applied researchers might usefully compare the character of knowledge in the social sciences and the character of school practitioners' knowledge and skills, in order to better understand connections and differences between the two realms. Federal research appropriations are now being cut with the same abandon that earlier marked their application, but in the midst of this the Secretary of Education asserted his belief that the dissemination of existing social science knowledge in education would greatly improve practice. There is no doubt that research has something to teach practice, and that researchers can learn much of value from practitioners. But in light of the record thus far, some closer scrutiny of how this learning might occur, and how it can usefully be encouraged, seems long overdue.

In the case of federal reform, matters are more complex. Some reform efforts, like Title I of the 1965 ESEA, or the enforcement of Title VI of the 1964 Civil Rights Act, represent major moral and political

commitments, and the forces at work in such programs are in many respects quite different from those considered in this essay. Here we have investigated a different species of federal reform--namely, programs which express no important national priorities, and for which the chief inspiration is not popular pressure, but administrators' or legislators' bright ideas about how to improve education. How might such efforts be best managed? One possibility--one of several, listed here only by way of example--is to conceive the federal role in such local reform as analogous to that of a foundation. The agency gives money to promising programs on the basis of the quality of local proposals. Some self-evaluation and financial accounting are required at the end. And the grantor hopes that things will turn out well in a respectable proportion of cases. Foundations have neither the resources nor the authority to require more, and in most cases, neither do federal education agencies. Such a posture would leave little room for the overblown aspirations of many federal officials, it would reduce administrative clutter, and would also improve federal-local relations. It might even encourage a more accurate picture of the federal role in certain sorts of local school reform.